

Université de Montréal

**Influencia interlingüística y transferencia morfosintáctica
en los aprendientes de español como lengua tercera (EL3)
en el País Vasco Norte (Francia)**

par Hugues Lacroix

Études hispaniques
Département de littératures et de langues du monde
Facultés des arts et des sciences

Thèse présentée
en vue de l'obtention du grade de Ph.D
en Littérature
option Études hispaniques

Août 2020

© Hugues Lacroix, 2020

Université de Montréal

Études hispaniques : Département de littératures et de langues du monde,
Faculté des arts et des sciences

Cette thèse intitulée

Influencia interlingüística y transferencia morfosintáctica en los aprendientes de español como lengua tercera (EL3) en el País Vasco Norte (Francia)

Présentée par
Hugues Lacroix

A été évalué(e) par un jury composé des personnes suivantes

Enrique Pato
Président-rapporteur

Anahí Alba de la Fuente
Directrice de recherche
(Université de Montréal)

Jean Casenave
Directeur de recherche
(Université Bordeaux-Montaigne)

Juana Liceras
Membre du jury

Ricardo Etxepare
Membre du jury

Maria Irene Moyna
Examinatrice externe

Résumé

Cette thèse porte sur l'influence interlinguistique au cours du processus d'acquisition de l'espagnol comme langue tierce (L3) chez des enfants bilingues français/basque. Plus précisément, nous nous intéressons à l'accord en genre dans le syntagme déterminant et à la formation de composés nominaux (N+N et déverbaux). Le français et l'espagnol partagent les mêmes caractéristiques concernant les deux phénomènes puisqu'ils possèdent chacun deux genres grammaticaux (masculin et féminin) et qu'ils forment leurs composés avec la tête à gauche. Par contre, en basque, il n'y a pas de genre grammatical et les composés se forment avec la tête à droite. Cette configuration des langues, ainsi que les différents profils de nos participants, nous permettent d'explorer l'influence interlinguistique dès les étapes initiales d'acquisition de l'espagnol L3 à travers l'étude des facteurs les plus courants dans le domaine de l'acquisition de L3, tels que la typologie, l'usage et l'ordre d'acquisition des deux premières langues. Dans l'ensemble, une influence du français (typologiquement et structurellement plus proche de l'espagnol) pourrait avoir un effet facilitateur, tandis que le basque aurait un effet inhibiteur.

Pour explorer ces possibilités d'influence, nous avons administré un test de description d'image écrite pour obtenir 1) des syntagmes formés par un déterminant, un nom et un adjectif de couleur (par exemple, « *la casa roja* », c'est-à-dire, « la maison rouge »); 2) des composés N+N (par exemple, « *hombre rana* », c'est-à-dire, « homme-grenouille »); et 3) des déverbaux (par exemple, « *abrebotellas* », c'est-à-dire, « ouvre-bouteille ») à 52 enfants bilingues (français/basque, âge : 10-11 ans) qui apprenaient l'espagnol comme L3 et à un groupe témoin de 21 locuteurs natifs d'espagnol du même âge. Les participants de L3 ont été divisés en trois groupes, selon leur langue dominante : L1 basque, L1 français et 2L1 basque/français. À partir des données collectées, nous avons analysé le niveau de maîtrise des structures cibles en fonction des facteurs les plus couramment étudiés en acquisition de L2. Concernant le genre grammatical, nous considérons le genre par défaut, le type d'accord (déterminant ou adjectif), la transparence de la terminaison (-a / -o) et le critère analogique. En ce qui a trait aux composés, nous nous intéressons à la directionnalité (tête initiale ou finale) et la productivité des deux types de composés. Plus spécifiquement pour les composés déverbaux, nous étudions également leurs caractéristiques formelles (flexion verbale et -s générique).

Conformément aux résultats de recherches antérieures sur le genre grammatical en espagnol, le taux d'erreur le plus élevé se produit avec le féminin, ce qui indique une stratégie consistant à utiliser le masculin comme genre par défaut. Nous avons également constaté que l'effet de la transparence de la terminaison était plus important chez les locuteurs ayant le basque comme langue dominante, qui ont acquis le français dans un contexte formel et qui ne l'utilisent pas quotidiennement à la maison. En revanche, le critère analogique semble avoir un impact plus important sur les locuteurs qui ont acquis le français à la maison et qui l'utilisent quotidiennement.

Concernant les composés, compte tenu des pourcentages élevés de composés avec la bonne directionnalité, nos résultats suggèrent un effet facilitateur du français pour tous les participants, quel que soit l'âge ou l'ordre d'acquisition de leur L1/L2.

En général, le français, la langue la plus proche, semble exercer une influence plus directe sur le processus d'acquisition de l'espagnol L3, mais le basque semble avoir une certaine influence sur les stratégies métalinguistiques mobilisées par nos participants.

Mots-clés : espagnol, acquisition de langues tierces, influence interlinguistique, proximité typologique, proximité structurelle, langue dominante, genre grammatical, composés N-N, composés déverbaux

Abstract

This study focuses on the issue of cross-linguistic influence (CLI) throughout the acquisition process of Spanish as an L3 by bilingual French/Basque children. Specifically, we are interested in gender concord within the Determiner Phrase (DP) and the formation of nominal compounds (N + N and deverbals). French and Spanish share the same characteristics regarding the two phenomena, as they display two grammatical genders (masculine and feminine), and compounds are head-initial. Conversely, there is no grammatical gender in Basque and compounds are head-final. This configuration of languages, together with the different profiles of our bilingual participants, allows us to explore cross-linguistic influence from the initial stages of acquisition of L3 Spanish through the most common factors in the field of L3 acquisition, namely typology, use and order of acquisition of the first two languages. Overall, an influence of French (typologically and structurally closer) would have a facilitating effect, while Basque would have an inhibitory one.

In order to explore these possibilities of influence, we administered a written image description task to elicit 1) phrases made up of a determinant, a noun and a color adjective (e.g., “*una casa roja*”, “a red house”), 2) N + N compounds (e.g., “*hombre rana*”, “frogman”) and 3) deverbal compounds (e.g., “*abrebotellas*”, “bottle opener”) to 52 bilingual children (French/Basque, age: 10–11 years) learning Spanish as an L3, and to a control group of 21 native speakers of Spanish of the same age. The L3 participants were divided into three groups, based on their dominant language: L1 Basque, L1 French, and 2L1 Basque/French. From the data collected, we analyzed the level of accuracy for the target structures in relation to the most commonly studied factors in L2 acquisition. Regarding grammatical gender, we considered the transparency of the ending (-a/-o), the default gender and the analogical criterion. Concerning the compounds, we studied directionality (head-initial vs. head-final) and the productivity of both types of compound. As for the deverbals, we also studied their formal characteristics (verbal inflection and generic -s).

In line with the results of previous research on Spanish grammatical gender, the highest error rates occur with feminine, which points to a strategy of using masculine as the default gender. We also found that the effect of the transparency of the ending was greater with Basque-dominant speakers, who acquired French later in a formal context and do not use it at home on a daily basis.

In contrast, the analogical criterion seems to be favoured by speakers who acquired French at home and use it every day.

As for the compounds, given the high percentages of compounds with the correct directionality, our results suggest a facilitating effect of French for all participants, regardless of age or order of acquisition of their L1/L2.

In general, French, the closer language, seems to exert a more direct influence on the process of acquisition of L3 Spanish, but Basque seems to have some influence in terms of the metalinguistic strategies used by our participants.

Keywords: Spanish, third language acquisition, cross-linguistic influence, typological proximity, structural proximity, dominant language, grammatical gender, deverbal compounds, N + N compounds

Resumen

La presente investigación se centra en la influencia interlingüística a lo largo del proceso de adquisición del español como L3 por parte de niños bilingües francés/euskera. Concretamente, nos interesamos por la concordancia de género dentro del sintagma determinante y la formación de compuestos nominales (N+N y deverbales). El francés y el español comparten las mismas características en cuanto a los dos fenómenos, al tener dos géneros gramaticales (masculino y femenino) y formar sus compuestos con el núcleo a la izquierda. En cambio, en euskera no hay género gramatical y los compuestos se forman con el núcleo a la derecha. Esta configuración de lenguas, junto con los diferentes perfiles de nuestros participantes bilingües, nos permiten explorar la influencia interlingüística desde las etapas iniciales de adquisición del español L3 a través de los factores más comunes en el campo de la adquisición de L3, como la tipología, el uso y el orden de adquisición de las dos primeras lenguas. De manera global, una influencia del francés (tipológica y estructuralmente más cercana y también la L2 de algunos participantes) podría tener un efecto facilitador, mientras que el euskera tendría un efecto inhibitorio.

Para explorar esas posibilidades de influencia, administramos una prueba escrita de descripción de imágenes para elicitación 1) sintagmas formados por un determinante, un sustantivo y un adjetivo de color (p. ej., “la casa roja”), 2) compuestos N+N (p. ej., “hombre rana”) y 3) deverbales (p. ej., “abrebotellas”) a 52 niños bilingües (francés-euskera, edad: 10-11 años) que aprendían español como L3 y a un grupo de control de 21 hablantes nativos de español de la misma edad. Los participantes de L3 fueron divididos en tres grupos, según su lengua dominante: L1 euskera, L1 francés, y 2L1 euskera/francés. A partir de los datos recogidos, analizamos el nivel de dominio de las estructuras metas en función de los factores más comúnmente estudiados en adquisición de L2. En cuanto al género gramatical, consideramos el género por defecto, el tipo de concordancia, la transparencia de la terminación (-a/-o) y el criterio analógico. Con respecto a los compuestos, estudiamos la direccionalidad (núcleo inicial vs. final) y la productividad de ambos tipos de compuesto. En cuanto a los deverbales, estudiamos además sus características formales (flexión verbal y -s genérica).

En línea con los resultados de investigaciones anteriores sobre el género gramatical en español, la tasa de error más alta ocurre con el femenino, lo cual apunta a una estrategia de uso del masculino como género por defecto. También encontramos que el efecto de la transparencia de la

terminación era mayor con los hablantes que adquirieron el francés en un contexto formal y que no lo usan en casa de manera cotidiana. En cambio, el criterio analógico parece tener un impacto mayor en los hablantes que adquirieron el francés en casa y que lo usan todos los días.

En cuanto a los compuestos, dados los altos porcentajes de producción de compuestos con la direccionalidad correcta, nuestros resultados sugieren un efecto facilitador del francés para todos los participantes, independientemente de la edad o del orden de adquisición de sus L1/L2.

De manera general, el francés, la lengua más cercana, parece ejercer una influencia más directa en el proceso de adquisición del español L3, pero el euskera parece tener cierta influencia en las estrategias metalingüísticas movilizadas por nuestros participantes.

Palabras clave: español, adquisición de lenguas terceras, influencia interlingüística, proximidad tipológica, proximidad estructural, lengua dominante, género gramatical, compuestos N+N, compuestos deverbales

Laburpena

Tesi hau frantses / euskaldun elebidunek gaztelania hirugarren hizkuntza gisa (H3) ikasterakoan gertzen den hizkuntza arteko eraginari buruzkoa da. Zehazkiago, generoarekiko komunztadura aztertzen dugu determinatzaile sintagman eta hitz elkartuetan (izen + izen eta aditz elkartuak). Frantsesak eta gaztelaniak ezaugarri berdinak dituzte bi fenomenoetarako dagokienez, bakoitzak bi genero gramatikal (maskulinoa eta femeninoa) baitituzte eta hitz elkartuek burua ezkerrean dute. Bestalde, euskaraz ez dago genero gramatikalik eta hitz elkartuek burua eskuinean dute. Hizkuntzen konfigurazio honek, baita gure parte-hartzaileen profil desberdinak ere, gaztelania H3 gisa jabetzerakoan hizkuntza arteko eragina aztertzeko aukera ematen digu, hasierako urratsetatik, H3 jabetzaren arloko faktore ohikoenen azterketaren bidez (hala nola hizkuntzen tipologia, erabilera eta eskuratzeko hurrenkera). Orokorrean, frantsesaren eraginak (tipologikoki eta egituraz gaztelaniatik hurbilago) eragin errazgarria izan dezake, eta euskarak, berriz, eragin inhibitzailea.

Eragin posible hauek aztertzeko, idatzizko irudi deskribapenaren proba pasatu dugu, 1) izen bat eta kolore adjektibo batez osatutako sintagma (adibidez, "*la casa roja*", hau da: "etxe gorria"); 2) izen + izen hitz elkartuak (adibidez, "*hombre rana*", hau da: "igel gizona"); eta 3) aditzez osatutako hitz elkartuak (adibidez, "*abrebotellas*", hau da, "botila irekitzekoa"). Gaztelania H3 gisa ikasten ari ziren 52 haur elebidunek (adina: 10-11 urte, frantsesa / euskara elebidun) eta gaztelania ama hizkuntza duten adin bereko 21 haur kontrolen talde batek parte hartu dute. Gaztelania H3ko partaideak hiru taldetan banatuak izan ziren, hizkuntza nagusiaren arabera: euskara H1, frantsesa H1 eta 2H1 euskara / frantsesa. Bildutako datuen oinarrian, xede-egituren ezagutza maila aztertu dugu bigarren hizkuntzaren jabetzaren aztertu ohi diren faktoreen arabera. Genero gramatikalari dagokionez, genero defektiboa, komunztadura mota (determinatzailea edo adjektiboa), bukaeraren gardentasuna (-a / -o) eta irizpide analogikoa kontuan hartzen ditugu. Hitz elkartuei dagokienez, bi hitz elkarte moten norabidea (hasierako edo amaierako burua) eta produktibitatea interesatzen zaizkigu. Berezi aditz elkartuentzat, haien ezaugarri formalak ere aztertzen ditugu (inflexioa eta -s generikoak).

Gaztelaniaz genero gramatikalari buruz egindako aurreko ikerketekin bat eginez, huts kopuru handiena femeninoarekin gertatzen da, maskulinoa genero defektibo gisa erabiltzeko estrategia erakutsiz. Gainera, ikusi da amaieraren gardentasunaren eragina handiagoa zela euskara

hizkuntza nagusi zuten eta frantsesa testuinguru formalean eskuratu eta etxean egunero erabiltzen ez duten hiztunengan. Aldiz, badirudi irizpide analogikoak eragin handiagoa duela etxean frantsesa ikasi duten eta egunero erabiltzen duten hiztunengan.

Hitz elkartuei dagokienez, norabide zuzena duten hitz elkartuen ehuneko altuak direla eta, gure emaitzek frantsesaren eragin errazgarria erakusten dute parte-hartzaile guztientzat, H1 / H2 jabetze adina edo hurrenkera edozein dela ere.

Orokorrean, frantsesak, hizkuntza hurbilena, gaztelania H3 jabetzeko prozesuan eragin zuzenagoa duela dirudi, baina badirudi euskarak nolabaiteko eragina duela gure parte hartzaileek mobilizatutako estrategia metalinguistikoetan.

Hitz gakoak: gaztelania, hirugarren hizkuntza jabekuntza, hizkuntza arteko eragina, hurbiltasun tipologikoa, hurbiltasun estrukturala, hizkuntza nagusia, genero gramatikala, izen-izen hitz elkartuak, aditz elkartuak

Índice

Résumé.....	i
Abstract	iii
Resumen.....	v
Laburpena.....	vii
Índice.....	ix
Lista de tablas	xiii
Lista de figuras	xiv
Lista de abreviaturas	xv
Lista de las abreviaturas utilizadas en las glosas	ii
Agradecimientos	iv
Introducción.....	1
CONTEXTOS SOCIOLINGÜÍSTICO	2
LOS FENÓMENOS ESTUDIADOS	4
<i>El género gramatical</i>	4
<i>La formación de los compuestos</i>	5
OBJETIVOS	5
ESTRUCTURA DEL TRABAJO	6
1 Adquisición de lenguas nativas y no nativas.....	8
1.1 CONCEPTOS CLAVE PARA DESCRIBIR AL INDIVIDUO BI-/MULTILINGÜE	9
1.1.1 <i>Momento de adquisición</i>	10
1.1.2 <i>Contexto de adquisición y tipos de conocimientos</i>	11
1.1.3 <i>La dominancia, el nivel de dominio, el uso y la cronología</i>	14
1.1.4 <i>Multicompetencia</i>	16
1.1.5 <i>Influencia interlingüística</i>	17
1.2 ADQUISICIÓN DE L1/L2	18
1.2.1 <i>L1</i>	18
1.2.1.1 El enfoque generativo	19
1.2.1.2 Perspectivas basadas en el uso	24
1.2.1.3 Conclusión parcial	29
1.2.2 <i>2L1</i>	30
1.2.3 <i>L2</i>	35
1.3 ADQUISICIÓN DE L3/LN	40
1.3.1 <i>L3</i>	40
1.3.2 <i>Nociones clave de AL3</i>	43
1.3.2.1 Edad, contexto y orden de adquisición	44
1.3.2.2 Conocimiento metalingüístico.....	45
1.3.2.3 Distancia tipológica y psicotipológica.....	46
1.3.2.4 Asociación de extranjería (<i>foreignness</i>) o el estatus de la(s) L2	48
1.4 MODELOS DE ADQUISICIÓN DE L3	49
1.4.1 <i>Hipótesis del factor de la L1</i>	49
1.4.2 <i>Hipótesis del factor del estatus de la L2</i>	50
1.4.3 <i>Modelo de la primacía tipológica</i>	51
1.4.4 <i>Modelo de mejoramiento cumulativo</i>	52

1.4.5	<i>Modelo de proximidad lingüística</i>	53
1.4.6	<i>Modelo de competición</i>	54
1.4.7	<i>Un modelo basado en el uso</i>	56
1.4.8	<i>Hipótesis generales con respecto a los modelos de AL3</i>	59
1.5	CONCLUSIONES	60
2	El género gramatical	62
2.1	NOCIONES PRELIMINARES	63
2.1.1	<i>Hacia una definición del género</i>	63
2.1.2	<i>Asignación y concordancia: dos caras de la misma moneda</i>	65
2.2	EL GÉNERO GRAMATICAL EN ESPAÑOL, FRANCÉS Y EUSKERA	66
2.2.1	<i>Español</i>	67
2.2.1.1	Asignación	67
2.2.1.2	Concordancia	69
2.2.2	<i>Francés</i>	70
2.2.2.1	Asignación	71
2.2.2.2	Concordancia	72
2.2.3	<i>Euskera</i>	75
2.2.3.1	Asignación	76
2.2.3.2	Concordancia	77
2.2.4	<i>Síntesis de las comparaciones entre las lenguas del estudio</i>	78
2.3	ANÁLISIS MORFOSINTÁCTICO	79
2.4	ADQUISICIÓN Y PROCESAMIENTO DEL GÉNERO GRAMATICAL	82
2.4.1	<i>En L1</i>	82
2.4.2	<i>En 2L1</i>	87
2.4.3	<i>En L2</i>	98
2.4.4	<i>En L3</i>	105
2.5	SÍNTESIS PROVISIONAL	106
2.6	DISEÑO EXPERIMENTAL	107
2.6.1	<i>Hipótesis</i>	109
2.6.2	<i>Participantes</i>	111
2.6.3	<i>Instrumentos</i>	114
2.6.3.1	Antecedentes lingüísticos y extralingüísticos	114
2.6.3.2	Pruebas diagnósticas	115
2.6.3.3	Tarea de denominación de imágenes	116
2.6.4	<i>Protocolo de administración de las tareas</i>	119
2.7	RESULTADOS	120
2.7.1	<i>Francés L1/2L1/L2</i>	121
2.7.2	<i>Euskera L1/2L1/L2</i>	126
2.7.3	<i>Español L1</i>	128
2.7.4	<i>Español L3: las etapas iniciales</i>	129
2.7.5	<i>Español L3: perspectivas longitudinales</i>	135
2.8	DISCUSIÓN	138
2.8.1	<i>El papel de los indicios morfofonológicos</i>	141
2.8.2	<i>El papel del criterio analógico</i>	142
2.8.3	<i>Influencia interlingüística y conocimientos metalingüísticos</i>	144
3	La formación de los compuestos	145
3.1	NOCIONES PRELIMINARES	147
3.1.1	<i>Núcleo, endocentricidad y exocentricidad</i>	148
3.1.2	<i>Subordinación y coordinación</i>	151
3.1.3	<i>Direccionalidad</i>	153
3.1.4	<i>Atomicidad</i>	156

3.1.5	<i>Compuestos sintagmáticos y compuestos propios</i>	157
3.1.6	<i>Ortografía y patrones de acentuación</i>	157
3.1.7	<i>Productividad, creatividad y recursividad</i>	158
3.2	LA COMPOSICIÓN EN ESPAÑOL, FRANCÉS Y EUSKERA.....	161
3.2.1	<i>Español</i>	162
3.2.1.1	Compuestos N+N.....	162
3.2.1.2	Compuestos deverbales	170
3.2.2	<i>Francés</i>	173
3.2.2.1	Compuestos N+N.....	173
3.2.2.2	Compuestos deverbales	181
3.2.3	<i>Euskera</i>	184
3.2.3.1	Los N+N	184
3.2.3.2	Los deverbales.....	189
3.2.4	<i>Síntesis de la comparación de las tres lenguas</i>	193
3.3	ANÁLISIS MORFOSINTÁCTICO Y SEMÁNTICO	195
3.3.1	<i>Compuestos N+N</i>	196
3.3.1.1	El parámetro de la composición	196
3.3.1.2	El marcador de palabra y la restricción de los corchetes dobles.....	200
3.3.1.3	Consideraciones semánticas	202
3.3.2	<i>Compuestos deverbales</i>	209
3.3.2.1	Consideraciones semánticas	211
3.4	ADQUISICIÓN DE LOS MECANISMOS DE COMPOSICIÓN	215
3.4.1	<i>En L1</i>	215
3.4.2	<i>En 2L1</i>	218
3.4.3	<i>En L2</i>	220
3.4.4	<i>En L3</i>	226
3.5	SÍNTESIS PROVISIONAL	229
3.6	DISEÑO EXPERIMENTAL	231
3.6.1	<i>Hipótesis</i>	232
3.6.2	<i>Participantes</i>	237
3.6.3	<i>Instrumentos</i>	238
3.6.3.1	Tarea de denominación de imágenes.....	239
3.6.4	<i>Protocolo de administración de las tareas</i>	243
3.7	RESULTADOS.....	244
3.7.1	<i>Los N+N</i>	245
3.7.1.1	Euskera L1/2L1/L2	245
3.7.1.2	Francés L1/2L1/L2	254
3.7.1.3	Español L1.....	258
3.7.1.4	Español L3: las etapas iniciales	261
3.7.1.5	Español L3: perspectiva longitudinal	267
3.7.2	<i>Los deverbales</i>	269
3.7.2.1	Euskera L1/2L1/L2	269
3.7.2.2	Francés L1/2L1/L2	275
3.7.2.3	Español L1.....	279
3.7.2.4	Español L3: las etapas iniciales	285
3.7.2.5	Español L3: perspectiva longitudinal	294
3.8	DISCUSIÓN	297
3.8.1	<i>Las etapas iniciales en español L3</i>	297
3.8.2	<i>Evolución longitudinal en español L3</i>	305
3.8.3	<i>Los compuestos en las dos primeras lenguas (euskera y francés)</i>	306
3.8.3.1	Los N+N y los deverbales en euskera	306
3.8.3.2	Los deverbales en francés y el rasgo de animacidad	309
4	Discusión general y conclusiones	312
4.1	SÍNTESIS GENERAL	312

4.1.1	<i>Modelos de adquisición de L3</i>	312
4.1.2	<i>El género gramatical</i>	314
4.1.3	<i>Los compuestos</i>	315
4.1.4	<i>Comparación de los fenómenos</i>	315
4.2	INFLUENCIA INTERLINGÜÍSTICA	318
4.2.1	<i>El orden de adquisición</i>	318
4.2.2	<i>La proximidad tipológica o estructural</i>	319
4.2.3	<i>La lengua dominante</i>	322
4.2.4	<i>Estrategias metalingüísticas</i>	323
4.3	REFLEXIONES PARA LA ENSEÑANZA DE ELE	324
4.4	PRINCIPALES APORTACIONES	327
4.5	INVESTIGACIONES FUTURAS	329
Bibliografía		331
Anexo 1.	Cuestionario para los padres (versión francesa)	I
Anexo 2.	Mini-diccionario visual	VI
Anexo 3.	Lista de ítems para elicitación del género	VII
Anexo 4.	Lista de ítems para elicitación de los compuestos N+N	IX
Anexo 5.	Lista de ítems para elicitación de los compuestos de verbales	X
Anexo 6.	Modelo de la evolución longitudinal de la concordancia de género	XI
Anexo 7.	Modelo de la concordancia de género (recolección 2)	XII
Anexo 8.	Modelo de la concordancia de género (recolección 3)	XIII
Anexo 9.	Resultados por ítem en la concordancia determinante – sustantivo	XIV
Anexo 10.	Resultados por ítem en la concordancia sustantivo - adjetivo	XV
Anexo 11.	Certificado ético	XVI

Lista de tablas

Tabla 1.1. Principales etapas del desarrollo del lenguaje en un niño según Clark (2009, pp. 14-15).....	19
Tabla 1.2. Resumen de las principales características de los modelos de AL3.....	58
Tabla 2.1. Lista de determinantes en francés	74
Tabla 2.2. Ejemplos de sustantivos en euskera que tienen una forma diferente en masculino y en femenino, junto con sus equivalentes en español.....	77
Tabla 2.3. Resumen del perfil de los participantes según la(s) lengua(s) hablada(s) en casa.	113
Tabla 2.4. Repartición de los ítems según las variables del estudio.....	118
Tabla 2.5. Resultados del modelo regresión logística de efectos mixtos de la concordancia de género en francés	125
Tabla 2.6. Tasa de concordancia en francés (L1, L2 o 2L1) y en español (L1).....	129
Tabla 2.7. Resultados del modelo regresión logística de efectos mixtos de la concordancia de género en español L3	133
Tabla 3.1. Las funciones básicas de Jackendoff (2011, 2016) utilizadas en los compuestos N+N subordinativos del español a partir de los datos de Toquero-Pérez (2018, pp. 37-38).....	166
Tabla 3.2. Clasificación de los sustantivos de tendencia adjetival en francés (Arnaud, 2018, p. 21) con sus equivalentes en español.....	181
Tabla 3.3. Ejemplos de compuestos nominales en las tres lenguas del estudio.	195
Tabla 3.4. Fenómenos vinculados con el TCP y su parametrización según Snyder (2012, p. 287).....	197
Tabla 3.5. Frecuencia de los compuestos cuyo N2 es "láser" y sus variantes con preposiciones.	204
Tabla 3.6. Resumen del perfil de los participantes según la(s) lengua(s) hablada(s) en casa.	238
Tabla 3.7. Repartición de los ítems N+N según los criterios de género, animacidad y si están atestiguados o no.....	242
Tabla 3.8. Repartición de compuestos deverbales según su animacidad y si están atestiguados o no.....	242
Tabla 3.9. Productividad y direccionalidad de los N+N en las lenguas (casi) nativas de los participantes.	262
Tabla 3.10. Resultados del modelo regresión logística de efectos mixtos del tipo de respuesta para los estímulos N+N.....	265
Tabla 3.11. Respuestas esperadas y alternativas de los participantes, acompañadas de sus frecuencias en el CdEN.....	283
Tabla 3.12. Productividad y direccionalidad de los compuestos deverbales en las L1/L2 de los participantes.	285
Tabla 3.13. Resultados del modelo de regresión logística de efectos mixtos del tipo de respuesta esperada para los ítems de compuestos deverbales en español.	289
Tabla 3.14. Flexión de los deverbales para todas las respuestas de tipo V+N a un estímulo en infinitivo.	293
Tabla 3.15. Repartición de las respuestas con o sin -s genérica en función de la paráfrasis del estímulo.	293
Tabla 4.1. Resumen de las hipótesis de influencia interlingüística.	316
Tabla 4.2. Análisis de los sustantivos de la Unidad 1 del manual Aula Internacional 1	325

Lista de figuras

Figura 1. Mapa del territorio de habla del euskera, con las provincias del País Vasco Norte en azul.	2
Figura 1.1. Continuum de los contextos de adquisición, de lo más formal a lo más informal.....	13
Figura 1.2. Jerarquía para clasificar las lenguas adquiridas, traducida de Hammarberg (2010, p. 101).....	41
Figura 2.1. Ítems 11 y 18 de la prueba de denominación de imágenes en español.	117
Figura 2.2. Tasa de concordancia de género por grupo y condición en la prueba de control en francés. .	124
Figura 2.3. Tasa media de concordancia por grupo según las diferentes condiciones analizadas.	132
Figura 2.4. Evolución de la tasa media de concordancia por grupo según el género y el tipo de concordancia.....	136
Figura 2.5. Evolución de la tasa media de concordancia por grupo según la transparencia y la congruencia	137
Figura 3.1. Ítems 59 y 52 de la prueba de denominación de imágenes en español.	240
Figura 3.2. Repartición de los tipos de respuestas según los ítems N+N en euskera.	247
Figura 3.3. Repartición de la direccionalidad de las respuestas N+N en euskera según los ítems.....	248
Figura 3.4. Direccionalidad de los compuestos N+N en euskera, según los grupos lingüísticos.....	250
Figura 3.5. Repartición de los tipos de respuestas según los ítems N+N en francés.....	256
Figura 3.6. Repartición de las respuestas según los ítems N+N en español L1.	259
Figura 3.7. Distribución de las respuestas para los ítems N+N. Todos los grupos experimentales (EU, FR, FR/EU) aparecen juntos.	263
Figura 3.8. Repartición del tipo de respuesta según el ítem en español L3.....	264
Figura 3.9. Evolución de la productividad y la direccionalidad de los NN - Grupos experimentales (la escala en Y comienza a partir de 70 % para hacer más visibles las diferencias entre los grupos).	268
Figura 3.10. Repartición de las respuestas según los ítems de compuestos deverbales en euskera.	271
Figura 3.11. Repartición de las respuestas según la morfología de los deverbales, por grupo y por animacidad en euskera.....	272
Figura 3.12. Repartición de las respuestas según el ítem en francés.....	276
Figura 3.13. Repartición de las respuestas según la morfología de los deverbales, por grupo y estructura.	278
Figura 3.14. Repartición de los tipos de respuestas según los ítems de compuestos deverbales en español L1.	282
Figura 3.15. Repartición del tipo de respuesta según los ítems deverbales en español L1, ordenados por frecuencia de las alternativas en el CdEN (solo aparecen los ítems que tenían formas alternativas).	284
Figura 3.16. Distribución de las respuestas para los ítems de compuestos deverbales en español. Todos los grupos experimentales (EU, FR, FR/EU) aparecen juntos.....	288
Figura 3.17. Morfología del verbo de los compuestos deverbales (n=624) - Grupos experimentales.	290
Figura 3.18. Flexión verbal de los compuestos deverbales en español L3, según la paráfrasis del estímulo.	292
Figura 3.19. Evolución de los aciertos en cuanto a la -s genérica y la flexión verbal en los compuestos deverbales.....	296

Lista de abreviaturas

2L1	Dos lenguas primeras	L2	Lengua segunda
AL2	Adquisición de L2 (disciplina)	L3	Lengua tercera
AL3	Adquisición de L3 (disciplina)	LN	Lengua nativa
ASALE	Asociación de Academias de la Lengua Española	LNN	Lengua no nativa
CdEN	Corpus del Español: <i>NOW</i>	LPM	<i>Linguistic Proximity Model</i>
CEM	<i>Cumulative Enhancement Model</i>	MCER	Marco Común Europeo de Referencia
CM	<i>Competition Model</i>	NGLE	Nueva Gramática de la Lengua Española
DLE	Diccionario de la Lengua Española	NP	<i>Noun Phrase</i>
DP	<i>Determiner Phrase</i>	RAE	Real Academia Española
ELE	Español como Lengua Extranjera	SD	Sintagma Determinante
ERP	<i>Event-related Potential</i>	SN	Sintagma Nominal
GC	Gramática de Construcciones	TCP	<i>The Compounding Parameter</i>
L1	Lengua primera/materna	TPM	<i>Typological Primacy Model</i>

Lista de las abreviaturas utilizadas en las glosas

En la medida de lo posible, hemos utilizado las abreviaturas de *Leipzig Glossing Rules* (Comrie et al., 2004), pero, para el euskera, esta lista de abreviaturas resultó insuficiente, por lo que hemos recurrido a algunas abreviaturas de la *Standard Basque progressive grammar* (de Rijk, 2008, pp. 1019-1025). La siguiente lista muestra todas las abreviaturas utilizadas a lo largo de la tesis.

ABR.	Nombre traducido al español (<i>original en inglés</i>)	Fuente
1	Primera persona (<i>first person</i>)	(Comrie et al., 2004)
2	Segunda persona (<i>second person</i>)	(Comrie et al., 2004)
3	Tercera persona (<i>third person</i>)	(Comrie et al., 2004)
ADN	Adnominal (<i>adnominal</i>)	(de Rijk, 2008)
ALO	Alocutivo (<i>allocutive</i>)	(de Rijk, 2008)
AUX	Auxiliar (<i>auxiliary</i>)	(Comrie et al., 2004)
DET	Determinante (<i>determiner</i>)	(Comrie et al., 2004)
ERG	Ergativo (<i>ergative</i>)	(Comrie et al., 2004)
F	Femenino (<i>feminine</i>)	(Comrie et al., 2004)
GEN	Genetivo (<i>genitive</i>)	(Comrie et al., 2004)
INS	Instrumental (<i>instrumental</i>)	(Comrie et al., 2004)
M	Masculino (<i>masculine</i>)	(Comrie et al., 2004)
PL	Plural (<i>plural</i>)	(Comrie et al., 2004)
PTCP	Participio (<i>participle</i>)	(Comrie et al., 2004)

REL	Relativo (<i>relative</i>)	(Comrie et al., 2004)
SG	Singular (<i>singular</i>)	(Comrie et al., 2004)
SOC	Asociativo (<i>sociative</i>)	(de Rijk, 2008)

Agradecimientos

Si bien esta tesis doctoral es el fruto de un trabajo personal importante, no hubiera sido posible sin el apoyo que muchas personas me brindaron en los últimos cinco años, por lo que quiero expresar mi gratitud general a todas las personas que me ayudaron de una manera u otra, tanto en la Université de Montréal, como en la Université Bordeaux-Montaigne y en el centro IKER-UMR-5478.

De manera más precisa, quiero agradecer a mis directores, Anahí Alba de la Fuente y Jean Casenave. Anahí, gracias por confiar en mi proyecto desde el inicio, por los intercambios de ideas que me ayudaron a delimitarlo mejor y por las incontables horas de reuniones que tuvimos tanto en Montreal como a distancia durante mis varios meses (y años) de estancia en el País Vasco Norte. La distancia no limitó de ninguna manera tu gran contribución a mi tesis. Jon, gracias por haberme recibido en el centro IKER, por tus consejos, por tu interés por mi proyecto también casi desde el inicio y tu ayuda siempre que la necesité.

Gracias a los *Fonds de recherche du Québec – Société et culture* (FRQSC) por la beca doctoral que representó para mí una ayuda financiera esencial. Gracias también a la *Maison internationale*, al *Département de littératures et de langues du monde*, a las ESP, a la FAÉCUM, a la *École Doctorale Montaigne Humanités* y a *Eusko Ikaskuntza* por las ayudas financieras adicionales que me permitieron asistir a congresos y financiar mi proyecto.

Al Centro de lenguas y al departamento por las cargas de curso que me permitieron adquirir una experiencia de enseñanza valiosa a lo largo del doctorado. Un agradecimiento especial a Danielle Vaillancourt, que ha sido mucho más que una coordinadora para mí. Gracias a los profesores del departamento que me acompañaron a lo largo de esos numerosos años: Javier Rubiera, Enrique Pato, Juan Carlos Godenzi, Ana Belén Martín Sevillano, James Cisneros y también Catherine Poupeney Hart, hoy jubilada, pero que supo contagiarme desde mi primer año de licenciatura con su pasión por los estudios hispánicos. Gracias también a Estela Bartol Martín que pasó unos pocos años en el departamento, justo lo suficiente para trasmitirme su pasión por la enseñanza del español como lengua extranjera.

A los profesores que, sin compartir mucho tiempo conmigo, me dieron consejos en las conferencias, coloquios, simposios u otros eventos científicos y me ayudaron a crecer como

investigador. Pienso en particular en Raquel Fernández Fuertes, Juana Licerias, Elena Valenzuela, Dalila Ayoun, Jason Rothman, Jennifer Cabrelli Amaro y Roy Lyster, entre otros muchos. Pienso también en otros profesores que se dieron la pena de contestar a mis preguntas y dudas por correo electrónico como Kevin McManus, Kim Tracy-Ventura, Lourdes Ortega, Peter Ecke y María Irene Moyna, entre otros.

A mi “comité de suivi” en la Université Bordeaux-Montaigne: Valérie Joubert, Argia Olçomendy y Marie Pourquoié.

A todos los participantes y a sus padres por su participación voluntaria a mi estudio. Gracias a todo el personal de los colegios que me ayudaron en la logística y me facilitaron el acceso a sus alumnos para la recogida de datos, particularmente a Céline, Eñaut, Iratxe, Odile, Iñaki y Xan en el País Vasco Norte y a Pedro y Susana, en España.

A todos los colegas y amigos del departamento en Montreal, particularmente a Maura y César que siempre estuvieron allí para ayudarme. Gracias también a Gabrielle, Vanessa y Laura.

A los colegas del *Colloque VocUM* con quienes emprendí una gran aventura de divulgación científica hace ya 6 años, particularmente Julien, Carlianne, Marie-Christine y Eve-Marie. Extiendo estos agradecimientos a todos los organizadores de *VocUM* desde el inicio y hasta ahora. Espero que este evento que fundamos durante mi maestría siga en pie de manera perenne para ayudar a futuros estudiantes de posgrado de la misma manera que me ayudo a mí.

A la ilustradora Edith Lebel cuyos bonitos dibujos me ayudaron a crear una prueba más amena para los niños.

A mis amigos del País Vasco Norte que me animaron a continuar y que me ayudaron de varias formas a lo largo de los últimos años: Marie, Maia, Xabi, Kaiet, Kiki, Anaiz, Olivier y Noël.

A mis amigos Gabriel, Etienne, David, Gabrielle, Simon-Pierre, Nicolas, Solène y Fernando, a los que no he visto mucho en los últimos años, pero que contribuyeron a formar la persona que soy ahora.

Si bien este proyecto de doctorado empezó hace cinco años, no hubiera sido posible escribir esta tesis en español si, hace 12 años, no hubiera conocido a todas estas personas que cruzaron mi camino durante mis dos años de andanzas por las carreteras mexicanas. Estos amigos me transmitieron su lengua y suscitaron en mí el deseo de emprender una carrera en estudios hispánicos

a mi vuelta a Canadá. Sin duda olvido a varios, pero allí va la lista, del noroeste al sureste de México: Mario, Cristián, Alejandro, Erman, Dianela, Jorge, Brenda, Shalo, Hector, Fernando, Katya, Adriana, Laurence, Rodrigo, Mónica, Violeta, Gama, Mario, Lorena, Luis, Manuel, Cristina, Pamela, Sergio, Eder, Mayte, Beatriz, Edwin, Nathalie, Miriam, Gabriela y Clara.

Gracias a mis padres y a mis hermanos que me animaron a continuar sin necesariamente entender lo que hacía durante todos estos últimos años de estudios.

Por último, un agradecimiento muy cariñoso a Amande, que se cruzó en mi camino justo antes de mi regreso a los estudios hace un poco más de 10 años, que me apoyó a lo largo de todos estos años y que espera desesperadamente que termine mi doctorado para poder pasar a otra etapa.
Milesker.

“Gizonen lana jakintza dugu: ezaguz aldatzea”
(Mikel Laboa)

Introducción

En un mundo en el que más de la mitad de la población habla dos o más lenguas (García Mayo, 2012, p. 130), la adquisición de lenguas terceras (AL3) está cada vez más presente a nivel global (Cenoz, 2013) y ha despertado un creciente interés en la comunidad científica internacional (Cenoz et al., 2001). A pesar de ello, pocos estudios se han enfocado en el proceso de adquisición del español como L3 en áreas bilingües como la región del País Vasco Norte (al suroeste de Francia), donde el francés y el euskera suelen ser las dos lenguas previamente adquiridas.¹

En la comunidad científica existe cierto consenso con respecto a la influencia que las lenguas previamente adquiridas pueden tener en la adquisición de una L3. Sin embargo, continúa el debate con respecto al peso de dicha influencia y a los factores que la determinan (De Angelis, 2007). Dentro de los factores más importantes, se suelen considerar el orden de adquisición (cronológico), el nivel de competencia en las lenguas previamente adquiridas y la distancia tipológica y estructural (las similitudes entre las lenguas). En los últimos años, han surgido varios modelos que se basan en los factores antes mencionados para predecir el proceso de adquisición de una L3. Ya que existe evidencia a favor y en contra de cada modelo teniendo en cuenta distintas combinaciones de lenguas, se hace cada vez más necesario comprobar las predicciones de cada modelo con combinaciones de lenguas menos comunes (Rothman, 2015, p. 189). Ante esta problemática, nuestra tesis estudia el trío francés-euskera-español en el País Vasco Norte.

¹ En nuestro trabajo, usamos el término País Vasco Norte para referirnos a la zona de habla del euskera situada en el suroeste de Francia, más precisamente, las provincias históricas de Labort, Baja Navarra y Sola. Si bien el nombre de País Vasco francés es de uso más común en español, optamos por seguir la terminología de la *VI Encuesta Sociolingüística* que se llevó a cabo sobre todo el territorio de habla del euskera (Eusko Jaurlaritzza - Gobierno Vasco, 2019). Además, “País Vasco Norte” corresponde a la traducción de los nombres más usados en el territorio mismo, es decir, “*Pays basque nord*” en francés (“País Vasco Norte”) e “*Iparralde*” en euskera (“Lado norte”), ambos en referencia a su ubicación al norte de los Pirineos.

Contexto sociolingüístico

La mayoría de los estudios sobre la adquisición de L3 que implican el euskera como una de las lenguas previamente adquiridas (García Mayo & Olaizola, 2011; García Mayo & Slabakova, 2015; González Alonso et al., 2016; Gutiérrez Mangado & Martínez-Adrián, 2018; Gutiérrez-Mangado & Martínez-Adrián, 2018; Imaz Agirre & García Mayo, 2013, 2014; entre otros) se han realizado en la Comunidad Autónoma del País Vasco (España). Por esta razón, consideramos oportuno proporcionar algunas precisiones sobre el territorio en el que se llevó a cabo el presente estudio: el País Vasco Norte, en Francia (ver Figura 1).

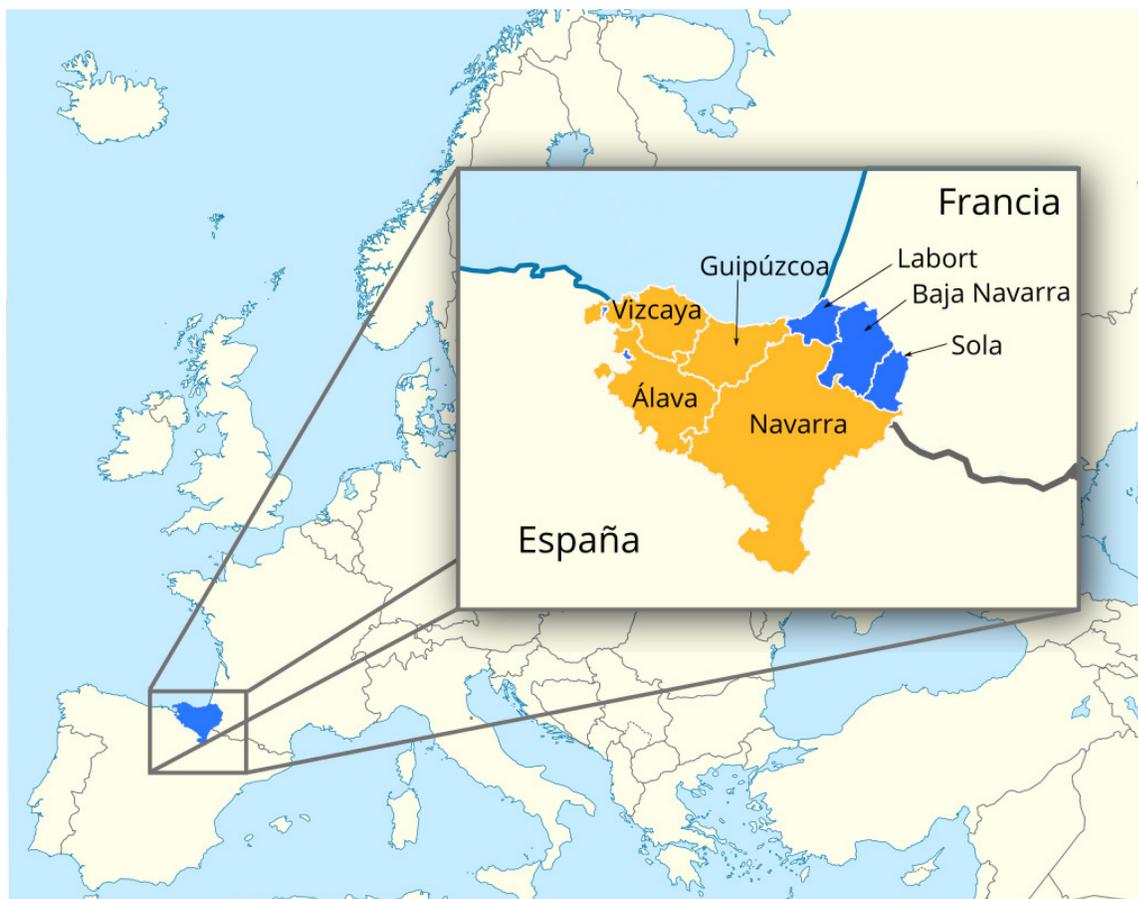


Figura 1. Mapa del territorio de habla del euskera, con las provincias del País Vasco Norte en azul.

Fuente: [Location of the Basque Country with labelled provinces in Europe Map](#),

por [Gabriel Trisca](#) / [CC BY-SA 3.0](#) (traducción y modificaciones menores nuestras)

En esta región, la lengua mayoritaria es el francés y una de las lenguas minoritarias es el euskera. Dado que el euskera no tiene ningún estatus oficial en la legislación francesa, la lengua de enseñanza de las escuelas públicas en Francia es siempre la única lengua oficial de este país: el francés. Sin embargo, las escuelas de los participantes de nuestro estudio son escuelas privadas bajo contrato de asociación y, por lo tanto, la lengua de enseñanza puede ser otra, en este caso, el euskera.² De manera general, los niños que asisten a estas escuelas, conocidas como *ikastola*, son hablantes nativos de francés, euskera o bilingües francés/euskera y, sea cual sea su lengua materna, tienen un alto nivel de competencia en ambas lenguas, como veremos más adelante cuando describamos a los participantes más detalladamente.

Según la *VI encuesta sociolingüística* (Eusko Jaurlaritz - Gobierno Vasco, 2019), en 2016, el territorio del euskera contaba con 3 131 464 habitantes, de los cuales solo 299 129 vivían en el País Vasco Norte. En todo el territorio, 751 527 habitantes (28,4 %) se consideraban como bilingües y 434 370 (16,4 %) se consideraban como bilingües pasivos.³ En el País Vasco Norte, había 51 197 (20,5 %) bilingües y 23 299 (9,3 %) bilingües pasivos. En esta región, dada la proximidad con la frontera española, el español es una lengua extranjera importante.

Nuestros participantes provenían de dos centros educativos distintos ubicados en las provincias de Baja Navarra y de Labort. Al inicio del estudio, todos estaban en sexto grado del

² Las escuelas privadas bajo contrato de asociación con el estado francés se asemejan a las escuelas concertadas españolas. Resultan de un acuerdo o concierto entre el estado y una entidad privada o asociativa y reciben financiamiento público. En el caso de las *ikastola*, la entidad vinculada por contrato con el estado se llama Seaska, una asociación gestionada por los padres de los alumnos. Si bien la *ikastola*, entendida como una escuela cuya lengua vehicular es el euskera, es común a todo el territorio vascohablante (Euskadi, Navarra y País Vasco Norte), los detalles de su implementación (pública, privada, privada concertada) varían de una zona a otra según la legislación del país o de la comunidad autónoma.

³ Todos los porcentajes fueron calculados a partir del número de habitantes de 16 años o más. La lengua común de todos los bilingües del territorio es el euskera, pero la otra lengua puede ser, o bien el español en la zona ubicada en España, o bien el francés en la zona ubicada en Francia. Las encuestas sociolingüísticas del País Vasco siempre hablan de bilingües (francés-euskera o español-euskera), ya que no existen hablantes monolingües de euskera. Los bilingües son definidos en la encuesta como los que son capaces de “hablar y entender bien o bastante bien el euskera”, mientras que los bilingües pasivos no son capaces de “hablar bien o bastante bien el euskera”, pero su conocimiento del euskera les permite al menos “entender y hablar algo el euskera” o “entender bien o bastante bien el euskera” (p. 209).

sistema de educación francés (11-12 años) y habían escogido el español como lengua viva 1 (LV1). En el sistema educativo francés, la LV1 corresponde a la primera lengua extranjera enseñada. En los colegios en los que se llevó a cabo nuestro estudio, los estudiantes tenían que elegir entre el español y el inglés como LV1. Solo participaron en el estudio los que habían escogido el español como LV1 y no habían tenido contacto significativo previo con el español.⁴ Entonces, el español representaba su tercera lengua según el orden de adquisición. Este es el criterio que utilizaremos para definir la L3 en el marco de nuestro trabajo.

Los fenómenos estudiados

Hemos decidido centrarnos en dos fenómenos lingüísticos que permiten contrastar los diferentes pares de lenguas según sus semejanzas (español-francés) y diferencias (español-euskera): el género gramatical y la formación de los compuestos. Estos fenómenos nos permiten explorar la influencia interlingüística, ya que las dos primeras lenguas de nuestros participantes muestran configuraciones diferentes. De esta manera podremos responder a la siguiente pregunta general: ¿cuál de las lenguas previamente adquiridas puede ser considerada como la fuente de influencia en cuanto al fenómeno X? Dado que los dos fenómenos que estudiamos han recibido mucho interés en el campo de la Adquisición de lenguas segundas (AL2), contamos con numerosos trabajos que nos servirán de base para nuestro estudio (Franceschina, 2005; Lardiere & Schwartz, 1997; Liceras & Díaz, 2000b; Liceras & Klassen, 2019; Montrul et al., 2008; White et al., 2004; entre otros).

El género gramatical

El francés y el español presentan un sistema con dos géneros (masculino y femenino) y materializan la concordancia de género de manera equivalente tanto en el determinante como en el adjetivo. En cambio, el euskera no materializa este género gramatical. Los ejemplos en (1a-b), con el núcleo subrayado y los elementos que concuerdan destacados en negrita, ilustran la concordancia del determinante y del adjetivo con el sustantivo en francés y en español, y la ausencia de concordancia de género en euskera.

⁴ A partir de su quinto grado en el colegio (2017-2018), todos empezaron los cursos de inglés como LV2, pero no tenían cursos de inglés durante el primer año de la recogida de datos (2016-2017).

- (1) a. el libro blanco (en masculino)
 le livre blanc
liburu zuria
- b. la casa blanca (en femenino)
 la maison blanche
etxe zuria

La formación de los compuestos

Hemos escogido dos estructuras que también permiten observar semejanzas y diferencias entre las lenguas en cuanto a la formación de los compuestos. Así, como se muestra en los ejemplos (2a-b), los compuestos de tipo Nombre-Nombre (N+N) y los deverbales (V+N) son similares en francés y en español, tanto en los rasgos formales como en la productividad de estos: el núcleo se encuentra a la izquierda, la composición N+N es poco productiva y la composición V+N es productiva. El euskera, en cambio, muestra características divergentes: el núcleo se encuentra a la derecha y la composición N+N es productiva. Para una mejor comprensión de los ejemplos, el núcleo de los compuestos aparece subrayado:

- (2) a. hombre araña (N+N)
homme araignée
 armiarma gizon
- b. lavaplatos (deverbal)
lave-vaisselle
 ontzi-garbigailu

Objetivos

En vista del creciente interés por el estudio de la AL3 y del constante aumento de trabajos enfocados en la L3 desde el punto de vista de la gramática generativa, el presente estudio tiene como objetivos generales: 1) analizar y profundizar en los aspectos metodológicos de los estudios empíricos de AL3, con especial interés en las dificultades y desafíos metodológicos propios al estudio de la adquisición multilingüe y 2) contribuir al campo de la AL3 con datos provenientes de una combinación de lenguas poco estudiada: el trío francés-euskera-español.

De manera más específica, el presente proyecto busca profundizar en la cuestión de cómo influyen las lenguas previamente adquiridas en el proceso de adquisición del español como L3 en un contexto en el que el aprendiente posee una lengua tipológicamente próxima (el francés) y otra

sin parentesco tipológico (el euskera). Para ello, hemos estudiado los dos fenómenos lingüísticos que presentamos anteriormente (el género gramatical y la formación de los compuestos), con el fin de explorar cómo interactúan las lenguas entre sí y, si existe una influencia, observar si esta ocurre de manera global (en ambos fenómenos) o parcial (fenómeno por fenómeno). Además, dado que nuestro estudio es longitudinal, podremos observar cómo evolucionan los efectos iniciales de la influencia a lo largo del tiempo. En línea con los estudios anteriores en el campo de AL3, exploraremos la influencia potencial de las dos lenguas previamente adquiridas y evaluaremos el impacto de los factores más relevantes para la influencia interlingüística: el orden de adquisición y el nivel de dominio de la de las dos primeras lenguas; el estatus de la lengua dominante; el uso (contexto y frecuencia) de las dos primeras lenguas; la proximidad tipológica entre L1-L3 y L2-L3; y la proximidad estructural entre L1-L3 y L1-L3. Para cada uno de estos factores, nos proponemos responder a la pregunta: ¿cuál es la importancia del factor Y en el proceso de influencia del fenómeno X? Las respuestas nos ayudarán a determinar qué modelo de AL3 (ver capítulo 1, sección 1.4) coincide mejor con la influencia observada en el marco de nuestra tesis.

Dado que, con estos objetivos y preguntas de investigación, las hipótesis son numerosas y dependen a la vez de cada modelo de AL3 considerado y del fenómeno estudiado, hemos agrupado nuestras hipótesis por modelo al final del capítulo 1, a las cuales se añaden hipótesis más precisas al tratar de los fenómenos lingüísticos de manera más específica (capítulos 2, 3).

Estructura del trabajo

Esta tesis se organiza en 4 capítulos. Dedicaremos el primer capítulo a los fundamentos teóricos generales de nuestro trabajo, es decir, la adquisición de lenguas nativas y no nativas. Definiremos conceptos clave que nos servirán para describir al individuo bi- o multilingüe, antes de adentrarnos en las teorías de la adquisición de la lengua nativa, para posteriormente cubrir la adquisición bilingüe y la adquisición de una L3. Presentaremos nociones clave y modelos propios del campo de AL3 y plantearemos hipótesis generales en función de estos modelos. Los principales modelos de adquisición de L3 que consideraremos en el marco del presente trabajo se inscriben en el marco generativo, lo que, sin embargo, no nos impedirá explorar también propuestas más afines a los acercamientos basados en el uso (*Usage-based approaches*). Siguiendo a Rothman y Slababoka (2018) y a Zyzik (2017), consideramos que ambas perspectivas convergen en varios

aspectos y que pueden complementarse para cubrir diferentes aspectos de los fenómenos lingüísticos.

Los dos capítulos centrales estarán dedicados a los dos fenómenos lingüísticos estudiados: el capítulo 2 versará sobre el género gramatical y el capítulo 3 sobre la formación de los compuestos. Estos dos capítulos incluyen los planteamientos teóricos que subyacen a cada fenómeno, nuestra metodología, nuestros resultados, así como una discusión sobre estos. Las bases teóricas que presentaremos para cada uno de estos capítulos consisten en una descripción detallada del fenómeno en las tres lenguas del estudio, un análisis morfosintáctico y una revisión de los trabajos de adquisición en lenguas nativas y no nativas. Esto significa que, durante una gran parte de cada capítulo, nos distanciaremos del aspecto de la L3 para enfocarnos en los fenómenos lingüísticos en cuestión, lo que nos permitirá, posteriormente, explorar la adquisición del fenómeno en español como L3 con bases más sólidas.

En el capítulo 4, volveremos sobre los resultados de los dos capítulos centrales para aunarlos y compararlos. Retomaremos las predicciones de los modelos de AL3 para organizar la discusión de nuestros resultados bajo dicha óptica. También abordaremos las implicaciones que nuestra aportación podría tener en la enseñanza del español como tercera lengua. Para terminar, ofreceremos las conclusiones generales, así como algunas propuestas para investigaciones futuras en el campo de adquisición de L3 de manera general y también, de manera más precisa, sobre las lenguas implicadas en la presente tesis. Los anexos que se encuentran después de la bibliografía incluyen información más detallada sobre la recogida y el análisis de los datos que puede resultar de ayuda especialmente durante la lectura de los capítulos 2 y 3.

“Of course, until we have crystal clear evidence that is truly irrefutable, we will continue to agree to disagree on nontrivial points. But disagreement should not be a bottleneck to progress for the broader [Second Language Acquisition] field.”
(Rothman & Slabakova, 2018, p. 436)

1 Adquisición de lenguas nativas y no nativas

El término “adquisición” aparece a menudo opuesto al término “aprendizaje”, sobre todo desde el trabajo de Krashen (1981). En su trabajo, Krashen planteó la adquisición como un proceso subconsciente que permite el desarrollo de la competencia lingüística de manera natural en contraste con el aprendizaje, entendido como un proceso consciente donde el aprendiente recibe enseñanza formal –o estudia con la ayuda de ciertos materiales de forma autónoma– prestando atención a la forma y a las reglas de la lengua. Para aludir a la combinación de los dos procesos, ciertos autores han utilizado el hiperónimo “apropiación” (Porquier, 1994 entre otros). Usando este término, podemos destacar que en la apropiación de una lengua nativa (LN/L1) predomina el proceso de adquisición, mientras que, en la apropiación de una lengua no nativa (LNN), el aprendizaje entra en juego de manera más significativa. Dado que el aprendizaje está vinculado a un proceso consciente y a la enseñanza formal, los trabajos en el campo de la educación suelen insistir más en el término aprendizaje, mientras que, en los trabajos orientados hacia el hablante, se privilegia el término adquisición. Como nuestro trabajo se enmarca en el campo de la adquisición de lenguas segundas (AL2) y terceras (AL3), utilizaremos principalmente el término adquisición, de manera que, a veces, abarcará también nociones que ciertos autores vinculan al aprendizaje (por ejemplo, Bange, 2002; Lambelet, 2012). Sin embargo, a la hora de hacer referencia a trabajos ajenos que hacen una distinción clara entre los dos términos, mantendremos dicha distinción.

Por su parte, la frontera entre LN y LNN parece, *a priori*, bien delimitada. De manera general, se podría decir que la LN es la lengua que se adquiere a partir del nacimiento (en la niñez) y todas las demás lenguas son no nativas. Sin embargo, esta clasificación no siempre resulta tan nítida, ya que existe una gran variedad de situaciones que pueden difuminar las fronteras entre LN y LNN. Según el contraste infancia vs. edad adulta, un niño podría adquirir dos lenguas en su infancia y tener, por lo tanto, dos LN. Como veremos más adelante, dentro del periodo de la infancia, la edad de adquisición puede tener un impacto a muchos niveles (tipo de bilingüismo, nivel de competencia en cada lengua, etc.). En vista de esto, antes de adentrarnos en las teorías de

adquisición de lenguas, presentamos a continuación una serie de conceptos clave vinculados al bi-/multilingüismo. Tras ello, presentaremos diferentes posturas teóricas sobre los procesos de adquisición de LN y LNN y contrastaremos los diferentes contextos de adquisición (L1, 2L1, L2, L3/Ln). Concluiremos el capítulo con una reflexión sobre las predicciones generales de las teorías de adquisición de L3 en el marco de nuestro estudio.

1.1 Conceptos clave para describir al individuo bi-/multilingüe

El objetivo de esta sección consiste en establecer conceptos clave –y relativamente neutros– para describir con más precisión nuestro objeto de estudio: el individuo que habla más de una lengua. De manera general, cuando un individuo habla más de una lengua, es considerado multilingüe⁵; cuando habla exactamente dos lenguas, se considera que es bilingüe, tres, trilingüe y así sucesivamente. Este bi-, tri- o multilingüismo puede tomar varias formas según el momento y el contexto de adquisición, el uso, la dominancia y el estatus sociopolítico de las lenguas, entre otros factores. Dado que el foco del presente proyecto se enmarca en el campo de la AL2/AL3 y se orienta más hacia el hablante que hacia su sociedad, no profundizaremos en el tema del estatus

⁵ Si bien existe una diferencia entre los términos multilingüismo y plurilingüismo, varios autores (sobre todo del mundo anglosajón) usan el término multilingüismo (*multilingualism*) para referirse a ambos conceptos. Según la definición del Marco Común Europeo de Referencia para las Lenguas (MCERL), el multilingüismo es “el conocimiento de varias lenguas o la coexistencia de distintas lenguas en una sociedad determinada”, mientras que el plurilingüismo se refiere al uso de varias lenguas por un mismo individuo que “no guarda esas lenguas y culturas en compartimentos mentales estrictamente separados” (Consejo de Europa, 2002, p. 4). La versión actualizada de los descriptores del MCERL insiste en el mismo concepto de plurilingüismo: “the fundamental point is that plurilinguals have a *single*, inter-related, repertoire that they combine with their general competences and various strategies in order to accomplish tasks” (Council of Europe, 2018, p. 28). En cambio, los que usan el término “multilingüismo” para referirse a ambos conceptos o “multilingüismo individual” para referirse a lo que el MCERL considera “plurilingüismo” suelen seguir un documento de la Comisión Europea que define el multilingüismo como “the ability of societies, institutions, groups and individuals to engage, on a regular basis, with more than one language in their day-to-day lives” (European Commission, 2007, p. 6). Si bien entendemos la diferencia entre los dos términos, dado que nuestra investigación no se centra en el “multilingüismo en la sociedad” y para mantener la coherencia con gran parte de la literatura que citamos, utilizaremos el término “multilingüismo” para referirnos al “plurilingüismo” o “multilingüismo individual”.

sociopolítico de las lenguas, sino que nos concentraremos en los factores vinculados con el hablante.

1.1.1 Momento de adquisición

En cuanto al momento de la adquisición, se suele diferenciar entre diferentes edades: por una parte, la edad de adquisición (EdA; *Age of Acquisition*) o edad de inicio (EdI; *Age of Onset*) y, por otra, la edad de primera exposición (EdE; *Age of First Exposure*). En el primer caso (EdA/EdI, los dos términos se usan de manera intercambiable), se considera que esa edad corresponde a la primera exposición significativa a la lengua, ya sea por inmersión o mediante cursos intensivos. En cambio, la EdE puede vincularse al momento en que una persona empieza a aprender una lengua en un contexto formal pocas horas a la semana, cuando se realiza una breve visita a un país donde se habla la lengua en cuestión, o cualquier otro contexto donde la exposición a la lengua no es tan importante (Birdsong, 2006).

En lo que respecta al bilingüismo infantil, ya que los niños empiezan a usar la lengua de manera más o menos fluida a partir de los tres años (Clark, 2009, p. 14), se suele diferenciar el bi-/multilingüismo que ocurre antes de los tres años (simultáneo) del que ocurre después (sucesivo o secuencial) (Montrul, 2013), pero no hay consenso al respecto y la frontera entre bilingüismo simultáneo y secuencial no está tan clara, como relata Hammarberg (2018) :

- (1.1) Researchers differ as to the timing of this process. Paradis (2008) places the end of the L1 acquisition period at about age five. In investigating child SLA, McLaughlin (1984) chooses the third birthday as a practical cutoff point between first and second language acquisition. Meisel (2011), reviewing research on neural maturation and age, finds that changes in the direction from L1-type to L2-type processing seem to take place at least up to age seven, with certain crucial changes around age four. Hyltenstam and Abrahamsson (2003), investigating age of onset and ultimate attainment, report effects that appear to indicate that the child's capacity for L1-type acquisition starts to decrease successively already shortly after birth. There is thus reason to interpret the L1/L2 distinction cautiously in the case of languages with an onset within this transition period. On the other hand, when dealing with a language acquired from early infancy or one learned in late childhood, adolescence or adulthood, the L1/L2 distinction will be clear-cut. (p.134)

De manera conservadora, nos quedamos con el límite de los 3 años basado en Clark (2009) y McLaughlin (1984) como EdA y concluimos que las lenguas adquiridas por un bi-/multilingüe simultáneo antes de esa edad son sus “lenguas nativas” (volveremos sobre el proceso de adquisición de dos lenguas nativas o 2L1 en la sección 1.2.2). Por el contrario, un hablante que

adquiere varias lenguas de manera secuencial –una antes de los tres años, otra(s) después– tendría una sola lengua nativa y otra(s) lengua(s) no-nativa(s). En cuanto a una lengua adquirida en el periodo de transición mencionado por Hammarberg (entre los 3 y 7 años), veremos a continuación qué otros elementos nos podrían ayudar a discernir entre lengua nativa y no nativa.

Más allá del tipo de bilingüismo (simultáneo/secuencial), ciertos autores consideran que la EdA condiciona el nivel alcanzado en una LNN, cuestión que retomaremos en la sección sobre la adquisición de L2 (1.2.3).

1.1.2 Contexto de adquisición y tipos de conocimientos

En cuanto al contexto, se suele hablar de adquisición formal/institucional en oposición a informal/natural, según si las lenguas se adquieren en la escuela o fuera de ella, respectivamente (Grosjean, 2010, pp. 25-26). De manera general, las LN se adquieren en contexto natural y las LNN en contexto formal. Aunque menos frecuente, es posible también encontrar casos de adquisición no nativa en la edad adulta sin enseñanza formal (inmersión sin cursos, por ejemplo), pero, fuera de lo anecdótico, los estudios que muestran casos de adquisición completa de una LNN sin ninguna intervención formal son, a nuestro conocimiento, prácticamente inexistentes (el estudio de caso de Schmidt, 1984 forma parte de los pocos casos documentados y no se puede decir que se trate, propiamente hablando, de una adquisición completa).⁶ Sin embargo, existen algunos estudios sobre la adquisición sin enseñanza formal que muestran cierto éxito en las etapas iniciales de adquisición, tanto del léxico (Gullberg et al., 2010; Shoemaker & Rast, 2013) como de la gramática (S. E. Carroll & Widjaja, 2013).

Volviendo a los casos más clásicos de LN en contexto natural vs. LNN en contexto formal, se considera que el contexto de adquisición tiene un impacto sobre el tipo de conocimiento que se desarrolla. El contexto formal suele ir acompañado de más conocimientos explícitos sobre la lengua adquirida en tal contexto, mientras que, en el contexto natural, prima el conocimiento implícito (Bley-Vroman, 1989; Hulstijn, 2005; Hulstijn & De Graaff, 1994; Krashen, 1981;

⁶ Volveremos sobre el término “completo” para caracterizar la adquisición de una lengua en la sección sobre la adquisición bilingüe 1.2.2 y los hablantes de herencia. De momento, en oposición a la adquisición completa, podríamos decir que los aprendientes desarrollan una “variedad básica” de la lengua en términos de Klein y Purdue (1997), es decir, una lengua que carece de ciertos rasgos morfológicos de la lengua meta.

Paradis, 2004; entre otros). Si bien los autores recién citados difieren en cuanto a los detalles de sus respectivas propuestas, existe cierto consenso alrededor del predominio de los conocimientos implícitos en una LN y explícitos en una LNN. En cuanto a los detalles, por ejemplo, Krashen (1981) considera que la L1 se adquiere principalmente de manera implícita y que, si bien los conocimientos explícitos son necesarios en el proceso de adquisición de L2, solo desempeñan la función de monitor para que el hablante luego adquiera conocimientos implícitos en esa lengua. Por su parte, desde la perspectiva de las neurociencias del lenguaje, Paradis (2004, 2008) considera que existe un mecanismo distinto para cada lengua: la adquisición de competencias lingüísticas implícitas para la L1 y el aprendizaje de conocimientos metalingüísticos explícitos para la L2. Según Paradis, la L1 se adquiere de manera incidental, se almacena de manera implícita, se usa de manera automática y es sostenida por la memoria procedural, mientras que la L2 se aprende y se controla de manera consciente, se almacena de manera explícita y se apoya en la memoria declarativa. Paradis añade que puede ocurrir cierta “automatización” de los conocimientos explícitos en la L2 con el uso, es decir, que el hablante puede desarrollar conocimientos implícitos equivalentes que eliminarán la necesidad de recurrir a los conocimientos explícitos.

Volviendo a la relación entre las dicotomías natural/formal e implícito/explicito, cabe aclarar que la relación no es tan directa como podría presuponerse a partir del párrafo anterior. Si bien el contexto formal fomenta el desarrollo de conocimientos explícitos, también puede fomentar las competencias implícitas. Además, la oposición natural/formal debería ser considerada más como un continuum que como una dicotomía propiamente dicha. Collentine y Freed (2004) mencionan tres contextos vinculados a la enseñanza formal que van desde lo más formal a lo más natural. En primer lugar, encontramos la enseñanza en el aula que, a pesar de tender hacia los métodos más comunicativos, sigue teniendo un fuerte énfasis en las estructuras gramaticales. En segundo lugar, los autores hablan de las inmersiones locales donde los aprendientes suelen tener la misma LN, pero tienen que comunicarse en la LNN de inmersión, durante un tiempo determinado (un campo de verano, por ejemplo), tanto dentro como fuera del aula. Si bien las situaciones de comunicación más naturales pueden surgir fuera del aula, el aspecto formal de este contexto sigue siendo predominante. En tercer lugar, los autores mencionan los programas de estudios en un país donde se habla la lengua meta (*study abroad*), en los cuales los aprendientes suelen tener clases formales, pero tienen acceso a muestras auténticas de la lengua fuera del aula (pp. 155-156). Si bien en esos tres contextos el componente formal conserva un lugar importante,

se puede observar que el último se acerca más al contexto natural. Además de los tres contextos mencionados por Collentine y Freed (2004), cabe añadir otro donde la lengua meta se usa como medio de instrucción (*Content and Language Integrated Learning*; CLIL). Este contexto se puede encontrar cuando la lengua de instrucción es una lengua extranjera (p. ej. el inglés en España) o en casos en que se usa la lengua minoritaria de una región bilingüe como lengua de enseñanza (País Vasco, Irlanda, Cataluña, etc.). En el último caso, algunos alumnos de una clase pueden tener la lengua meta como L1, por lo que pueden surgir situaciones de comunicación más naturales en el aula. Además, a pesar de ser minoritaria, la lengua meta se puede utilizar fuera del aula en la vida cotidiana, como en el caso de los programas de *study abroad* mencionados por Collentine y Freed (2004). La Figura 1.1 retoma y reordena los diferentes contextos que acabamos de describir.

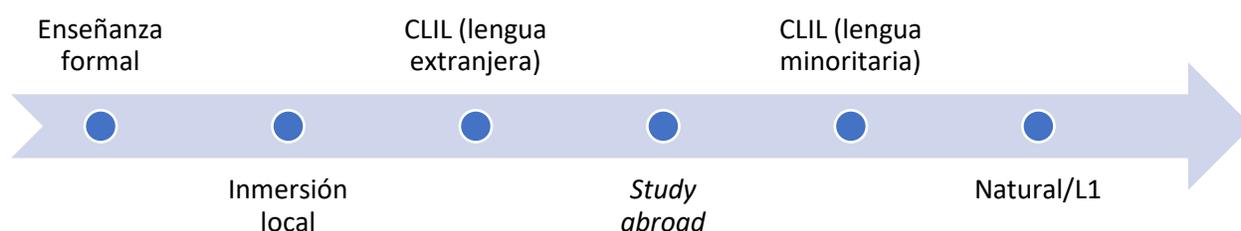


Figura 1.1. Continuum de los contextos de adquisición, de lo más formal a lo más informal.

Si bien la Figura 1.1 no representa todas las situaciones posibles, sirve para ilustrar algunos de los grados existentes entre la enseñanza formal y la adquisición en contexto natural. De hecho, se podría representar el continuum de manera multidimensional, tomando en cuenta el método de enseñanza privilegiado, la función de la lengua en el aula (lengua de instrucción o curso de lengua), la posibilidad de comunicarse con nativos dentro y fuera del aula, etc.

Dicho esto, se podría decir que, de manera general, los bilingües simultáneos adquieren sus dos primeras lenguas de manera esencialmente implícita mientras que los bilingües secuenciales tienen más probabilidades de adquirir su segunda lengua de manera explícita, si esta adquisición ocurre en un contexto formal. Sin embargo, dada la gran variedad de contextos de adquisición de la L2 para los bilingües secuenciales, es imprescindible dar una serie de precisiones necesarias a la hora de describir el nivel de bilingüismo de un hablante dado. En la sección 1.3.2, que se centra en los factores de influencia en el proceso de adquisición de L3, volveremos sobre un constructo relacionado con los conocimientos explícitos: la consciencia metalingüística.

1.1.3 La dominancia, el nivel de dominio, el uso y la cronología

La noción de dominancia es otra variable importante que entra en juego a la hora de caracterizar al bi-/multilingüe. Se trata de una noción multifacética; hasta la actualidad no existe consenso sobre la manera de concebir y medir dicho constructo (Silva-Corvalán & Treffers-Daller, 2016). De manera intuitiva, se considera que la lengua dominante de un hablante es la lengua en la cual tiene el mayor dominio y mayor fluidez.⁷ Sin embargo, la noción de nivel de dominio es, a su vez, un constructo multidimensional que abarca la sintaxis, la morfología, la semántica, la pragmática, la fonología, la fonética, y el vocabulario y resulta difícil establecer qué componente se debe medir y cuál no (Treffers-Daller, 2019). Además, la complejidad relativa de cada lengua no se refleja siempre de la misma manera en todos los componentes (Hulstijn, 2012), entonces, la adaptación de una prueba para medir varias lenguas y tener resultados comparables puede resultar difícil. A veces se utilizan pruebas de vocabulario, ya que sus resultados suelen estar correlacionados con el nivel de dominio en los demás componentes de la lengua (Milton, 2009). Una alternativa a esta medición del nivel de dominio de una lengua consiste en considerar el uso relativo de cada lengua de un hablante, entendido tanto en función de la exposición como de la producción, de acuerdo con el cual se considera que la lengua más usada es la lengua dominante. Esta medida en función del uso –a menudo correlacionada con otras medidas del nivel de dominio– está cobrando cada vez más importancia y ha sido utilizada en varios estudios recientes sobre la dominancia de las lenguas en el bilingüismo infantil (Bedore et al., 2012; Morgia, 2016; Unsworth, 2016).

Por otra parte, el uso, el nivel de dominio y la dominancia pueden no corresponder con otra variable a menudo utilizada para describir a los bi-/multilingües: el orden cronológico de adquisición. Es decir, un bilingüe que usa poco su LN/L1 bien podría no ser dominante y tener

⁷ El uso de la expresión “nivel de dominio” (que hemos tomado del MCERL y que se traduce por “niveau de maitrise” en francés o “proficiency level” en inglés) puede resultar confuso para el lector en contraste con el término “dominancia”. “Nivel de dominio” se refiere a lo que se puede medir concretamente mediante unas pruebas dadas y que puede corresponder a una nota numérica o una escala de un marco (por ejemplo, 90 % o B1), en otras palabras, una medida independiente de las demás lenguas del hablante. Por su parte, “dominancia” hace referencia a la relación entre dos o más lenguas, de las que la más “fuerte” se consideraría como la lengua dominante.

poca fluidez en esta lengua y, en cambio, tener un mayor dominio de su LNN/L2.⁸ Es el caso, por ejemplo, de inmigrantes que llevan varios años viviendo y trabajando en su LNN. A este respecto, en la última década, el caso de los hablantes de herencia como objeto de estudio (en la mayoría de los casos, hijos de inmigrantes que han adquirido una lengua en casa diferente de la lengua mayoritaria de la sociedad en la que viven) ha fomentado la consolidación de un campo de investigación propio (Brinton et al., 2008). En el caso de los hablantes de herencia, a menudo, el orden cronológico de adquisición no coincide con el orden de dominancia, ya que la lengua que han adquirido durante su infancia no suele utilizarse fuera del ámbito del ámbito familiar.⁹

En el transcurso de la vida de cada individuo bi-/multilingüe, sus lenguas tienden a evolucionar y a cambiar conforme a las diferentes variables que se han mencionado hasta ahora. Por lo tanto, algunos investigadores (Grosjean & Li, 2013; entre otros) prefieren usar etiquetas como La, Lb, Lc, etc. para referirse a las lenguas de un hablante, ya que estas no presuponen, *a priori*, ningún orden cronológico o jerárquico entre las lenguas. Sin embargo, como apunta De Angelis (2007), existen diferencias (que presentaremos más en detalle en la sección 1.3) entre una L2 y otra lengua adquirida después cronológicamente. A pesar de ello, se sigue utilizando a menudo la etiqueta L2 para referirse a cualquier lengua no nativa el campo de la AL2 y, hasta muy recientemente, el joven campo del multilingüismo (como rama de los estudios de AL2) carecía de una terminología más precisa para ilustrar esas diferencias. Las etiquetas L3, L4, L5 (orden cronológico o de dominancia) o Ln y LX (lenguas no nativas o lenguas adicionales después de la segunda) son algunas de las nuevas etiquetas que se han usado para diferenciar las lenguas del hablante multilingüe.

En el marco de este estudio, se usará “lengua nativa” (LN) para referirse de manera general a una (L1) o varias lenguas (2L1) adquiridas antes de los tres años en un contexto natural y “lengua no-nativa” (LNN) para todas las demás lenguas. Cuando sea necesario especificar el orden de

⁸ Como explica Grosjean (2010, pp. 20-21), frente a la idea tradicional de que el bilingüe era aquel hablante que podía expresarse a la perfección en sus dos lenguas, ‘sin acento’ y de manera equilibrada, el bi/multilingüe no tiene que ser ni equilibrado ni perfecto, sino que tiene que conocer más de una lengua y poder usarlas en diferentes contextos.

⁹ Volveremos a hablar de los hablantes de herencia más en detalle en la sección 1.2.2 sobre la adquisición bilingüe simultánea.

adquisición, utilizaremos las etiquetas L2 y L3, la L3 entendida como la lengua adquirida después de otras dos lenguas, ya sean estas nativas o no nativas, como en Cenoz (2003a).

1.1.4 Multicompetencia

La multicompetencia es otro concepto relevante que ayuda a definir al individuo bi-/multilingüe, en este caso, en alusión al estado complejo y dinámico de la mente. El término, acuñado por Cook (1991), se basa en el hecho de que, en comparación con los monolingües, los bi-/multilingües tienen competencias en más de una lengua, lo cual tiene implicaciones a varios niveles. Cook define la multicompetencia como un estado dinámico de la mente de los hablantes multilingües que implica la interacción entre las diversas representaciones mentales de las lenguas conocidas. Desde este punto de vista, resulta evidente que la L1 y la L2 de un bilingüe no pueden ser exactamente equivalentes a esas dos lenguas en dos hablantes monolingües. Para ilustrar su concepto, Cook (1992) cita estudios en los cuales la L1 se ve influida por la L2 y viceversa. Por ejemplo, en el área de la fonética, estudios sobre el inicio de sonoridad (VOT, del inglés *Voice Onset Time*) muestran cómo la L2, en este caso el inglés, influye en la pronunciación y la percepción de ciertas consonantes de la L1 (en francés L1: Flege, 1987; en español L1: Nathan, 1987). Esta influencia no resulta sorprendente, dado que la multicompetencia presupone que todas las lenguas conocidas por un hablante están siempre activadas en su mente. El nivel de activación de cada lengua puede variar según el momento o las circunstancias, pero no parece haber desactivación completa de una lengua en ningún momento. Si bien en los años noventa esta noción de mente compuesta –en contraste con una mente donde todas las lenguas están claramente aisladas– llevaba un cierto carácter innovador, desde entonces, numerosos trabajos de corte psicolingüístico apoyan esta noción de activación simultánea de varias lenguas a la vez (de Bot, 2004; Dijkstra, 2003; Marian & Spivey, 2003; entre otros). Por ejemplo, basándose en trabajos empíricos sobre el procesamiento léxico en los multilingües, Dijkstra (2003) y de Bot (2004) presentan modelos que toman en cuenta la activación relativa de las palabras de todas las lenguas del hablante multilingüe. En esos modelos, ciertos factores contribuyen a inhibir ciertas palabras de una lengua no deseada en un momento dado y otros factores contribuyen a la mayor activación de las palabras de la lengua deseada, pero todas las lenguas están siempre activadas hasta cierto punto. En las dos últimas décadas, los progresos en neuroimagen han permitido comprobar que el hecho de tener que inhibir de manera constante una o más lenguas contribuye a un mayor desarrollo

de las funciones ejecutivas del cerebro, es decir, las funciones responsables de la inhibición, la atención y la memoria a corto plazo (Bialystok et al., 2012, p. 241). Si bien el concepto de multicompetencia tal como lo acuñó Cook en 1991 no indicaba de manera tan precisa cuáles eran las consecuencias concretas sobre el cerebro, las premisas del constructo contribuyeron a un cambio en la percepción de los bi-/multilingües. La reciente publicación de un volumen completo (Cook & Li, 2016) dedicado a la noción de multicompetencia muestra que sigue siendo de actualidad en el campo del multilingüismo.

1.1.5 Influencia interlingüística

Como hemos visto anteriormente, la mente del bi-/multilingüe es muy dinámica y está sujeta a interacciones entre las múltiples lenguas presentes. La noción de influencia interlingüística –bajo el término “interferencia”– aparece por primera vez de manera clara en los años cincuenta en los trabajos de Weinreich (1953/1979) y Lado (1957). En aquella época, se empleaba el término interferencia con connotaciones negativas para explicar desviaciones de la norma por parte de los bilingües y los aprendientes de L2, una tendencia que se mantuvo en los trabajos de lingüística contrastiva hasta finales de los años setenta. Desde entonces, investigadores como Kellerman y Sharwood-Smith (1986) han ampliado el uso del constructo y lo han denominado, de manera más abarcadora y más neutra, “influencia interlingüística” (*Cross Linguistic Influence; CLI*, en inglés). Según Kellerman y Sharwood-Smith, esta expresión cubre no solo lo que Weinreich llama interferencia, sino también las estrategias de evasión y los préstamos.¹⁰ Desde entonces, se siguen usando los términos transferencia e influencia interlingüística (y, en menor medida, interferencia) de manera casi intercambiable, con el significado amplio introducido por Kellerman y Sharwood-Smith (1986). Si bien, como veremos en la sección 1.2.3, varios estudios han demostrado que el proceso de adquisición de L2 no se reduce a la influencia interlingüística, la noción sigue siendo central en el campo de AL2. La gran cantidad de trabajos sobre la influencia interlingüística (Gass, 1988; Kellerman, 1983, 1995; Kellerman & Sharwood Smith, 1986; Odlin, 1989, 2005, 2008; Selinker, 1983; entre otros) demuestra el interés en este tema por parte de los investigadores en este campo. En las últimas décadas, este concepto también ha sido explorado en el campo de AL3

¹⁰ Siguiendo a autores como Álvarez Benito y Iñigo Mora (2005) y Moreno (2010), hemos adoptado la expresión “estrategias de evasión” para referirnos a lo que Kellerman y Sharwood-Smith (1986) llaman “*avoidance strategies*”.

como un elemento clave para explicar parte del proceso de adquisición de las L3 (Alba de la Fuente & Lacroix, 2015; Cenoz et al., 2001; M. Ortega, 2008; entre otros).¹¹

La influencia interlingüística es un concepto multifacético de gran abarque que puede afectar a todos los componentes de la lengua, ya sea la semántica, la sintaxis, la morfología, la fonología, la fonética o la ortografía (Odlin, 2008). Además, la aplicación de esta noción no se limita al individuo, sino que pueden observarse fenómenos de influencia interlingüística a nivel de la sociedad. Concretamente, del mismo modo que la L1 de un hablante puede ejercer cierta influencia sobre su L2, la L2 sobre la L3, y viceversa, la lengua mayoritaria de una sociedad puede ejercer una influencia sobre una lengua minoritaria, y viceversa.¹²

Tras este breve panorama de algunos conceptos clave sobre los bi-/multilingües, a continuación, expondremos las bases teóricas del proceso de adquisición de una sola lengua nativa (L1) para luego comparar este proceso con los de adquisición de 2L1, L2 y L3.

1.2 Adquisición de L1/L2

1.2.1 L1

Todos los niños de desarrollo típico pueden adquirir una LN, sea cual sea, si están expuestos de manera suficiente a la lengua en cuestión (*input*). A pesar de estar expuestos a muestras de lengua incompletas, o incluso con errores, los niños logran adquirir un sistema lingüístico coherente, pasando por etapas similares alrededor de las mismas edades (la Tabla 1.1 resume las grandes etapas de desarrollo que exploraremos más en detalle a continuación). En los distintos marcos teóricos de la lingüística, la afirmación previa no se debate, sino que es un hecho observable empíricamente. Sin embargo, las bases subyacentes a ese proceso de adquisición pueden variar según el enfoque teórico que se adopte.

¹¹ Volveremos sobre el tema de la L3 con más detalle en la sección 1.3.2.

¹² Si bien los efectos de la L2 o la L3 hacia la L1 y de la lengua minoritaria hacia la mayoritaria han sido menos estudiados, trabajos como los de Restorick Elordi (2012) y Alba de la Fuente, Cruz Enríquez y Lacroix (2018) muestran que este tipo de influencia también es posible.

Tabla 1.1. Principales etapas del desarrollo del lenguaje en un niño según Clark (2009, pp. 14-15)

Etapa	Edad aproximada
Balbuceo	7-10 meses
Palabras aisladas	13-19 meses
Combinación de palabras	Alrededor de los 24 meses
Refinamiento del lenguaje	Hasta los 6 años
Dominio de las construcciones complejas	10-12 años

Desde los años 60, la gramática generativa, impulsada por Chomsky (1965, 1957/2002), se ha impuesto como el acercamiento predominante y ha suscitado la producción de una miríada de estudios, sobre una multitud de lenguas y fenómenos, tanto en LN como en LNN. Además del generativismo, a lo largo de las últimas décadas han ido surgiendo diferentes propuestas, como, por ejemplo, los acercamientos basados en el uso, con sus correspondientes alternativas sobre el proceso de adquisición del lenguaje. Las diferencias teóricas han provocado vivos debates entre generativistas y no generativistas, aunque, como veremos a continuación las diferentes posturas tienen varios elementos en común (Rothman & Slabakova, 2018). Si bien las justificaciones sobre qué acercamiento teórico es más acertado para estudiar la adquisición del lenguaje quedan fuera del alcance de la presente tesis, a lo largo de nuestros capítulos, exploraremos las contribuciones que ambos acercamientos proporcionan a los fenómenos estudiados. Las dos subsecciones siguientes presentan, de manera general, los rasgos principales de estos dos enfoques.¹³

1.2.1.1 El enfoque generativo

En el marco de la lingüística generativa, la facultad del lenguaje depende de un sistema mental innato que permite al hablante generar una posibilidad infinita de oraciones, es decir, incluso oraciones que nunca ha oído. En el momento de su aparición, la innovadora propuesta de

¹³ Con el objetivo principal de debatir sobre los argumentos a favor y en contra de cada enfoque, varios investigadores han dedicado tesis (Preston, 2015), artículos (Gregg, 2003; O’Grady, 2008b, 2008a, 2010; Rothman & Slabakova, 2018) y libros (O’Grady, 1997, 2005) al tema, por lo que, además de la información proporcionada en las subsecciones 1.2.1.1 y 1.2.1.2, remitimos el lector a las obras citadas para más detalles.

Chomsky (1959) se plantea en fuerte oposición a la tesis conductista de Skinner (1957), según la cual los niños adquieren el lenguaje mediante los principios de asociación e inducción, imitando y formulando generalizaciones a partir de lo que oyen. Según Chomsky, hay principios demasiado abstractos en el lenguaje para que se puedan aprender mediante simples mecanismos de asociación e inducción. Concretamente, el argumento clave en contra de Skinner –y a favor de un sistema innato del lenguaje– gira en torno al llamado problema lógico de la adquisición del lenguaje o de la pobreza del estímulo (*Poverty of Stimulus*, término que aparece por primera vez en Chomsky, 1980).

Ese problema parte del hecho de que los niños, a pesar de recibir un *input* a veces reducido e imperfecto, logran adquirir la lengua exitosamente. Según el argumento de la pobreza del estímulo, ciertas particularidades de las lenguas no siempre están presentes en el *input* que reciben los niños, lo que vuelve imposible la adquisición inductiva a partir de este. Además, los niños solo reciben muestras de lo que se permite en la lengua (evidencia positiva) y raras veces reciben instrucciones sobre lo que una lengua no permite (evidencia negativa). A pesar de eso, los hablantes nativos tienen una intuición clara sobre las estructuras que no se permiten en su lengua, aun sin haberlas observado antes.

Para explicar la adquisición exitosa de la lengua a pesar de ese *input* relativamente pobre, los generativistas postulan la existencia de un dispositivo de adquisición del lenguaje (DAL) y una gramática universal (GU; Chomsky, 1965). El primer término se refiere a la capacidad innata de todos los seres humanos para adquirir el lenguaje. Existe cierto consenso en cuanto a su existencia, pero hay debates sobre su forma y su contenido (Liceras & Díaz, 2000, p. 5). A su vez, la GU conforma el estado inicial de la gramática de cualquier lengua, es decir, un inventario de principios, categorías y rasgos universales subyacentes a todas las lenguas. Desde los inicios de la lingüística generativa, lo que se considera como parte de la GU ha ido evolucionando, partiendo de unas simples reglas transformacionales en las primeras versiones de la teoría (Chomsky, 1965, 1957/2002), pasando por un inventario extenso de principios y parámetros en los años ochenta (Chomsky, 1980, 1981) para llegar a un inventario mínimo en las últimas décadas (Chomsky, 1995, 2000, 2008). De hecho, en los últimos años algunos generativistas han argumentado a favor de un solo principio que formaría parte de la GU: el principio de recursividad (Hauser et al., 2002).

A partir de la GU y del *input* que recibe, el niño configura su lengua-I – donde la *I* hace referencia al carácter individual, interno e intensional de la lengua– para que esta corresponda a la lengua a la cual está expuesto.¹⁴ De ese modo, la GU proporciona un punto de partida común para iniciar el proceso de adquisición del lenguaje y también contribuye a explicar que todos los niños pasen por etapas de adquisición similares, en edad comparable, sea cual sea su lengua. El *input* y las propiedades de la lengua meta pueden tener un impacto directo en algunas etapas de desarrollo de la lengua-I, pero el resultado final es que, alrededor de los cuatro o cinco años, todos los niños de desarrollo típico llegan a dominar su lengua materna (Montrul, 2013).

Según el modelo de Principios y Parámetros (Chomsky, 1981), la adquisición de la sintaxis se hace mediante la fijación de parámetros de manera binaria. Si bien los principios son universales (los mismos para todas las lenguas), la configuración específica de los parámetros es la responsable de la variación entre las lenguas. Por ejemplo, según el principio de proyección ampliado (*Extended Projection Principle, EPP*, Chomsky, 1982), todos los predicados deben tener un sujeto, pero no todas las lenguas cumplen con ese principio de la misma manera. Aquí entra en juego el parámetro del sujeto nulo o *pro-drop*, para el que el español tiene el valor [+sujeto nulo] y el francés [-sujeto nulo], lo que permite explicar que en español se puedan producir oraciones sin sujeto explícito y que no exista esta posibilidad en francés. Dentro de la teoría de los Principios y parámetros, el parámetro del sujeto nulo no está solamente relacionado con la presencia explícita de los sujetos en las oraciones, sino que se considera como un macro-parámetro que se asocia con una serie de propiedades de la lengua (como la flexión verbal rica/pobre, la posición pre- o posverbal del sujeto, entre otras propiedades, según Jaeggli, 1982; Rizzi, 1982). De esta manera, en función del *input* que reciba, el niño fijará el parámetro en su valor positivo si el *input*

¹⁴ La intensionalidad de la lengua-I se refiere al hecho de que se trata de un sistema regido por reglas y que su contenido no está definido de manera extensiva y redundante. El término *I-Language* aparece por primera vez en Chomsky (1986), así como su contrapunto del *E-Language*, el lenguaje externalizado. Para referirse a la lengua-I, también se ha usado el concepto de competencia (*competence*) en oposición a actuación (*performance*) (Chomsky, 1965). Si la competencia corresponde a la representación interna de la lengua de un hablante, el conocimiento inconsciente e intuitivo, la actuación corresponde a la externalización de esa lengua mediante la producción (o la comprensión). En el proceso de externalización de la lengua, entran en juego varios factores y, por lo tanto, la actuación no es siempre un reflejo fiel de la competencia. La lingüística generativa ha tomado como objeto de estudio la competencia/lengua-I, es decir, la representación mental de una lengua específica en un individuo dado.

corresponde a una lengua de sujeto nulo o negativo en el caso contrario. En el modelo de Principios y Parámetros, la noción de macro-parámetro tenía un alto poder explicativo, ya que permitía exponer cómo, a partir de un *input* limitado, los niños podían adquirir varias propiedades de una vez.

Con la llegada del Programa Minimalista (Chomsky, 1995 y las aportaciones posteriores de 2001, 2007, 2008), llega un cambio en la formulación de los parámetros. Los macro-parámetros, es decir, la agrupación de varias propiedades de una lengua en un grupo como en el caso del parámetro del sujeto nulo, dejan de formar parte de la GU y la variación interlingüística se explica más bien mediante lo que Baker (2008) llama la Conjetura Borer-Chomsky: “all parameters of variation are attributable to differences in the features of particular items (e.g., the functional heads) in the lexicon” (p. 144). Según esta afirmación, los rasgos (*features*) se adquieren en el léxico y éstos mismos son responsables de la parametrización de una lengua-I. En este contexto, Licerias (2010) explica que, si bien todos los seres humanos tienen el mismo inventario de rasgos posibles, no todas las lenguas activan los mismos rasgos. En otras palabras, todos los rasgos forman parte de la GU, pero según el *input* recibido, un hablante solo asociará ciertos rasgos a ciertas categorías funcionales y léxicas para llegar a la configuración de su lengua-I.¹⁵

Si bien la lingüística generativa siempre ha asumido una separación entre la sintaxis y el lexicón, este último acercamiento sitúa la mayor parte de la variación interlingüística en el lexicón, es decir, fuera del módulo puramente sintáctico. Dicho de otra manera, las reglas formales de la sintaxis son, en esencia, las mismas para todas las lenguas y son los rasgos que se aplican a ciertos elementos del lexicón los que permiten explicar la variación.

¹⁵ Se considera generalmente, en la gramática generativa, que los verbos (V), los nombres (N), los adjetivos (A) y las preposiciones (P) representan las categorías léxicas; son categorías de inventario abierto y las palabras que contienen son designadoras o denotativas. Por otra parte, las categorías funcionales son de inventario cerrado y contienen palabras consideradas como meramente “gramaticales”, palabras que solo cumplen una función en la oración sin aportar ningún significado adicional. Algunos ejemplos de esas categorías son los determinantes (D), los complementantes (C), las conjunciones (Conj) y los auxiliares (v) (Bosque & Gutiérrez-Rexach, 2009, pp. 108-113). A esas categorías funcionales directamente asociadas a algunas palabras, la gramática generativa añade otras que solo se proyectan en el árbol sintáctico y que albergan rasgos funcionales como la categoría tiempo (T) (Adger, 2003, Capítulos 5-7).

Volviendo al *input* de manera más general, los sujetos nulos representan un ejemplo de cómo ese *input* puede servir para fijar los parámetros de una lengua. Los niños que adquieren inglés producen un 30-40 % de sujetos nulos, frente a cerca del 70 % producido por niños que aprenden lenguas como el español, el catalán o el italiano (Montrul, 2013). Por su parte, los niños que adquieren español producen artículos aproximadamente seis meses antes que los que aprenden inglés. Estos mismos niños hispanohablantes producen una mayoría de verbos con flexión de persona, número, tiempo y modo, mientras que los que aprenden inglés producen una mayoría de infinitivos (sin flexión), lo cual corresponde a la fase que se ha dado en llamar de “infinitivo raíz” (*root infinitive stage*; T. Harris & Wexler, 1996; Hoekstra & Hyams, 1998). Cada lengua tiene particularidades más o menos salientes –el sujeto nulo, el uso de artículos, la flexión verbal, etc.– que hacen que el sistema lingüístico se desarrolle de manera distinta de una lengua a otra (Montrul, 2013).

Como hemos visto en la presente subsección, si bien los generativistas otorgan un papel al *input* en el proceso de adquisición de la LN, ese papel se limita al efecto de desencadenante (*trigger*) de los parámetros o rasgos de la GU y de ninguna manera es suficiente por sí solo. En los últimos años, el generativismo ha sido criticado por investigadores de diversas ramas (psicolingüistas, neurolingüistas, entre otros) sobre las bases del papel menor otorgado al *input* y a la experiencia. Algunos autores llegaron a afirmar que ya se había comprobado la irrelevancia de la GU considerando los avances que se habían hecho en la lingüística cognitiva para responder al Problema lógico de la adquisición (Christiansen & Chater, 2015; Dąbrowska, 2015; Evans, 2014; Everett, 2016). Este punto ha sido cuestionado por parte de autores de corte generativista (Boxell, 2016; Mendivil-Giró, 2018; entre otros).¹⁶

Antes de pasar a la siguiente sección sobre los acercamientos basados en el uso, queremos subrayar la contribución innegable de la lingüística generativa en el desarrollo de un marco teórico formal para estudiar de manera detallada la sintaxis de varios fenómenos lingüísticos en varias lenguas. Por ello, tomamos los trabajos realizados dentro del marco generativista como punto de partida para nuestra investigación, a los que añadimos la exploración de propuestas alternativas con el objetivo de establecer un contacto entre ambas posturas. Somos conscientes de que este acercamiento es poco frecuente en el campo, pero nos parece interesante explorar las posibles

¹⁶ Los detalles del debate quedan fuera del alcance de la tesis y remitimos el lector a los autores citados.

semejanzas y contribuciones de cada perspectiva a los fenómenos de nuestro estudio. Pasamos, pues, a explorar los principios generales de las perspectivas basadas en el uso.

1.2.1.2 Perspectivas basadas en el uso

El título de la presente subsección aparece en plural, ya que se considera que no hay un solo acercamiento al tema, sino varias posturas que comparten, al menos, las dos características siguientes: 1) el *input* lingüístico es la primera fuente de la adquisición de lenguaje y 2) los mecanismos cognitivos que se movilizan para la adquisición del lenguaje son generales, no solo sirven para la adquisición del lenguaje. Dado el papel central que se otorga a la cognición en general en esta perspectiva –en oposición a la visión generativista de que el lenguaje está vinculado a un módulo propio–, se trata de una perspectiva que se inscribe en la corriente de la lingüística cognitiva.¹⁷ Si, por un lado, los generativistas se han basado en la pobreza del estímulo y la imposibilidad de adquirir la lengua mediante el *input* para argumentar a favor de un dispositivo innato del lenguaje, los cognitivistas, por su parte, se han dado en demostrar que la pobreza del estímulo no existe y que el lenguaje emerge solo a partir del *input*, las interacciones sociales y las capacidades cognitivas del ser humano (Tomasello, 2003). Desde las perspectivas basadas en el uso, no se argumenta en contra de la capacidad innata del ser humano para adquirir la lengua, sino que se afirma que esta capacidad innata no se debe a la GU sino los mecanismos cognitivos generales.

De todas las perspectivas basadas en el uso, tomaremos como punto de partida la propuesta de Tomasello (2003), que plantea un acercamiento que pretende explicar de manera holística tanto la adquisición de lenguas nativas como no nativas.¹⁸ A su vez, la propuesta de Tomasello se

¹⁷ Autores como García Mayo (2012) utilizan la expresión “*Cognitive Approaches*” para agrupar la perspectiva psicológica (que se asemeja a lo que presentamos aquí como “perspectivas basadas en el uso”) y la lingüística formal. En cambio, en lingüística cognitiva, no se suele considerar al generativismo como una rama del cognitivism (Ellis, 1999). En lingüística cognitiva, las propuestas son muy variadas de tal modo que un panorama completo no cabe en la presente tesis. Para más detalles sobre esta cuestión, remitimos al lector a Ellis (1999) y las fuentes que mencionamos más adelante en la página 25 sobre la gramática de construcciones y la gramática cognitiva.

¹⁸ De todas las propuestas cognitivistas, incluso el conexionismo, ninguna propone una teoría de alcance tan amplio como las perspectivas basadas en el uso como explica Tomasello (2003): “while connectionist models are an extremely valuable tool, and are definitely on the right course in many ways, in their current state they are not sufficient to

construye sobre las bases de la Gramática cognitiva y la Gramática de construcciones expuestas en los trabajos de Langacker (1987, 1987), Lakoff (1987, 1990), Fillmore (1988), Talmy (1988), Croft (1991) y Goldberg (1995). La clave de la adquisición del lenguaje según Tomasello (2003) se encuentra en dos aptitudes en los niños: (1) la habilidad de leer las intenciones comunicativas (*intention-reading skills*), asociada a la dimensión funcional del lenguaje y (2) la habilidad de encontrar patrones en las muestras de la lengua (*pattern-finding skills*), vinculada a la dimensión gramatical; dos habilidades que son esenciales para la adquisición del lenguaje, pero que no son exclusivas del lenguaje. La interacción de esas dos habilidades generales permite los cuatro procesos de adquisición que describe Tomasello (2009) y que reproducimos a continuación (1.2-1.5) :

- (1.2) *Lectura de intenciones y aprendizaje cultural*, que explica cómo los niños aprenden el vínculo forma-función, incluyendo todo desde las palabras solas hasta las construcciones complejas;
- (1.3) *Esquematización y analogía*, que explica cómo los niños pueden crear construcciones sintácticas abstractas (y papeles sintácticos como sujeto y objeto directo) a partir de los enunciados concretos que han oído;
- (1.4) *Atrincheramiento y preempción/priorización*, que explica cómo el niño constriñe sus abstracciones conforme a las que son convencionales en su comunidad lingüística; y
- (1.5) *Análisis distribucional basado en las funciones*, que explica cómo los niños forman categorías paradigmáticas de varios tipos de constituyentes lingüísticos (p. ej. los sustantivos y los verbos).¹⁹

Esos cuatro procesos, que van más allá del simple aprendizaje estadístico y lineal basado en la frecuencia, son considerados como suficientes para adquirir una lengua. En cuanto a la Pobreza del estímulo, Tomasello (2003, 2009) argumenta simplemente que no existe, ya que, a

account for child language acquisition because they do not include in any significant way information about communicative intention or function.” (pp. 145-146)

¹⁹ Traducción nuestra a partir del inglés:

- *Intention-Reading and Cultural Learning*, which account for how children learn conventional form–function pairings, including everything from words to complex constructions;
- *Schematization and Analogy*, which account for how children create abstract syntactic constructions (and syntactic roles such as subject and direct object) out of the concrete utterances they have heard;
- *Entrenchment and Preemption*, which account for how children constrain their abstractions to just those that are conventional in their linguistic community; and
- *Functionally Based Distributional Analysis*, which accounts for how children form paradigmatic categories of various kinds of linguistic constituents (e.g. nouns and verbs). (p. 89)

pesar de no tener una GU formal que permite constreñir las posibilidades de una lengua, los niños tienen esos otros mecanismos que permiten analizar las construcciones de la lengua y así discernir entre lo que es posible y lo que no es posible en su lengua. Dicho de otro modo, “[m]odern usage-based theorists are not behaviourists who believe the child works with unstructured linear strings, but rather they are cognitivists who believe in structure – just not of the purely formal kind.” (Tomasello, 2009, p. 89). De manera comparable, Bybee (2010) insiste en cinco procesos cognitivos generales –que no son propios del lenguaje– que interactúan entre sí para contribuir al desarrollo del lenguaje: (1) la categorización, que permite la organización ordenada de la información; (2) la fragmentación (*chunking*), que permite separar las unidades del *input*; (3) el almacenamiento de la memoria compleja (*rich memory storage*) que permite la acumulación de información sobre la experiencia, los contextos de uso, la información fonética de las palabras y las oraciones, (4) la analogía, que permite asociar los enunciados nuevos a los ya almacenados y, finalmente, (5) la asociación inter-modal, que permite el vínculo entre el significado y la forma. Esos cinco procesos –atestiguados también fuera del ámbito del lenguaje– culminan en la creación de las unidades y estructuras que constituyen el sistema lingüístico del hablante. A pesar de las diferencias en la nomenclatura, y la separación de los procesos, los acercamientos de Bybee (2010) y Tomasello (2009) son comparables en esencia.

Las estructuras a las que Tomasello y Bybee aluden son las llamadas construcciones, que representan la base del lenguaje. En el ámbito de las construcciones, no hay límites definidos entre la sintaxis y el léxico, sino que se contempla más bien un continuo que va desde el morfema hasta la oración completa, desde lo más concreto a lo más abstracto. Las construcciones pueden combinarse de modo que formen oraciones completas y llegar a representaciones formales comparables a las que se llega con una gramática de corte generativo. Por ejemplo, suponemos, por una parte, construcciones concretas que consisten en pronombres, sustantivos, determinantes, verbos y, por otra parte, otras más abstractas, por ejemplo, construcción nominal (determinante + sustantivo) y construcción transitiva (sujeto + verbo + objeto).²⁰A partir de la combinación de estas diferentes construcciones se puede formar una oración transitiva (por ejemplo, “el niño come un

²⁰ El objetivo de esta sección es presentar la noción de construcción y sus principios generales. El análisis de los detalles formales de las construcciones queda fuera del alcance de la tesis. Referimos al lector interesado a Tomasello (2003, Capítulo 5) para una descripción más detallada de esta propuesta teórica.

helado”) de manera comparable a la estructura que se formaría a partir del lexicón y de la operación “fusión” en la gramática generativa. Para formar oraciones más complejas, se pueden combinar estructuras abstractas del mismo modo que se proyectarían otros nodos en un árbol sintáctico de corte generativo. Así, podemos considerar otra construcción abstracta, que llamaremos “cláusula relativa epistémica”, consistente en “Sujeto + V epistémico + que + oración”. Combinando esta construcción con la anterior, podemos formar oraciones como “Yo creo que el niño come un helado”. Según los principios de la adquisición basada en el uso, todas estas estructuras abstractas proceden de las generalizaciones que se hacen a partir de construcciones más concretas. Por ejemplo, la estructura de la “cláusula relativa epistémica” procedería de una generalización a partir de *input* como “creo que ...”, “pienso que...”, “opino que”, etc. Gracias a los diferentes niveles de abstracción de las construcciones, se puede reproducir de manera exitosa el principio de recursión tan emblemático de la teoría generativa (Christiansen & Chater, 2015; Christiansen & MacDonald, 2009). Sin embargo, Tomasello (2003, pp. 99-100) defiende que las construcciones explican de manera más adecuada el proceso gradual de adquisición de la lengua por parte del niño. A partir del *input* que recibe, el niño empieza a adquirir ciertas construcciones sencillas (palabras concretas, aisladas) y luego aprende a combinarlas mediante construcciones cada vez más abstractas, dando lugar a un desarrollo lingüístico que va de lo más concreto a lo más abstracto, como el desarrollo de las demás capacidades cognitivas del niño.

Siguiendo a Tomasello (2003, pp. 108-122) y a otros autores como Gathercole et al. (1999) y Sebastián et al. (2004) para la adquisición del español, repasamos los diferentes tipos de construcciones con las que se topa el niño en las primeras etapas de su proceso de adquisición del lenguaje. La primera construcción del niño consiste en la holofrase, es decir, una palabra aislada que puede corresponder a una sola palabra adulta, por ejemplo, “beber”, “agua”, o a una serie de palabras que el niño analiza como una sola unidad, por ejemplo, “oabua”, que conlleva un significado propio para expresar una idea más compleja como “quiero beber agua”. Esta construcción mínima aparece en el repertorio del niño hacia los 12 meses de edad. Luego, hacia los 18 meses de edad, aparece el esquema pivote, es decir, una construcción de dos palabras que contiene una parte fija y un hueco rellenable, por ejemplo, “Más X” para pedir más de una cosa X, donde X podría ser “pan”, “zumo”, “postre” o cualquier otro elemento compatible con el esquema. En esta etapa, el orden de los componentes del esquema pivote no conlleva un cambio de significado, dado que esos esquemas no implican una estructura sintáctica propiamente dicha.

Después, hacia los dos años, el niño empieza a usar construcciones basadas en un ítem central. A menudo, el elemento central de estas construcciones es un verbo y, en estos casos, estas construcciones llevan el nombre de islas verbales. De manera similar al esquema pivote, las islas verbales tienen huecos rellenables, pero son más complejas, ya que los huecos corresponden a un orden específico y cada elemento cumple un papel sintáctico dentro de la construcción. Por ejemplo, un niño podría tener una construcción con el verbo querer del tipo “_____ quiere _____”, en el cual el primer hueco se llena con un sujeto y el segundo con cualquier elemento compatible, como un sustantivo o un verbo en infinitivo. Este proceso da lugar a enunciados como “nene quiere agua” o “nene quiere jugar”. En esta etapa, la morfología de manera general no cumple un papel importante en las construcciones del niño. Por ejemplo, un verbo conjugado en tercera persona del singular podría ser utilizado para referirse a una primera o segunda persona personal, otro verbo podría ocurrir siempre en primera persona del singular, otro en primera persona del plural, etc. Existe evidencia tanto para el español (Gathercole et al., 1999; Sebastián et al., 2004) como para otras lenguas (Tomasello, 2003, pp. 118-119) de que las islas verbales se adquieren una por una y de que, en esta etapa, el niño todavía no analiza completamente la composición de las islas y no es capaz de generalizar de una construcción a otra. A partir de los 3 años, el niño empieza a hacer generalizaciones de una construcción a otra y, poco a poco, las construcciones más frecuentes se vuelven más abstractas, lo que abre el paso a la adquisición de construcciones abstractas similares a las que emplean los hablantes adultos, como las construcciones transitivas, intransitivas, pasivas, imperativas, reflexivas, locativas, causativas, interrogativas, etc.

Esas construcciones abstractas emergen cuando el niño ha acumulado una cantidad suficiente de construcciones basadas en ítems y cuando es cognitivamente capaz de llevar a cabo lo que Tomasello llama el análisis distribucional basado en las funciones (*Functionally Based Distributional Analysis*). Dicho análisis permite al niño categorizar adecuadamente los componentes de las construcciones según las funciones que ocupan, establecer analogías entre elementos comparables y de allí llegar a establecer un inventario de construcciones más abstractas. En este proceso se establecen las categorías paradigmáticas que corresponderían, a grandes rasgos, a las categorías funcionales y léxicas de la gramática generativa. A partir de esta etapa, el niño empieza a hacer generalizaciones que resultan imprescindibles para el desarrollo de su lengua y a veces sobregeneraliza, lo que lo lleva a producir estructuras que se desvían de la norma adulta. Sin embargo, estas sobregeneralizaciones son relativamente poco frecuentes y, por lo tanto, cualquier

teoría de la adquisición del lenguaje debe ser capaz de explicar qué es lo que limita esas sobregeneralizaciones o, dicho de otro modo, lo que constriñe la gramática del niño. En el marco generativo, esta explicación procede de los principios de la GU, mientras que las propuestas basadas en el uso, como explica Tomasello, invocan tres procesos que concurren para bloquear las generalizaciones no deseadas: el atrincheramiento, la preempción y la formación de clases de verbos. El atrincheramiento es el proceso según el cual las construcciones más frecuentes se van asentando con más firmeza en la mente del hablante. Conforme a este principio, cuanto más atrincherada se vuelva una construcción, menos probable es su cambio y su generalización a otras funciones. De manera complementaria, el principio de preempción es él que bloquea la generalización de una construcción cuando el niño oye una construcción que compite con una posible generalización. Así, adaptando el ejemplo de Tomasello (2003, p. 178), podríamos decir que el hecho de oír una construcción como “Él hizo desaparecer el conejo” bloquea la posible generalización del uso transitivo del verbo desaparecer: “Él desapareció el conejo”.²¹ Finalmente, la formación de clases de verbos es lo que permite al niño asociar nuevos verbos con otros verbos que ya conoce según sus similitudes (funciones causativa, transitiva, ditransitiva, etc.) y así generalizar una construcción para una clase de verbo. Estos tres procesos, según Tomasello, son suficientes para constreñir la gramática del niño y explicar cómo este evita producir construcciones que se desvían de la lengua a la cual está expuesto.

1.2.1.3 Conclusión parcial

Los dos enfoques que acabamos de presentar explican desde diferentes perspectivas las distintas etapas del complejo proceso de adquisición del lenguaje por parte de los niños. Por una parte, el enfoque generativo considera que el niño cuenta con una capacidad innata como punto de partida. En contraste, los enfoques basados en el uso consideran que los procesos cognitivos generales son suficientes para explicar la adquisición del lenguaje. Según ambos enfoques, se considera que el lenguaje está formado por estructuras complejas que se pueden combinar para dar lugar a estructuras aún más complejas. En el enfoque generativo, estas estructuras son los sintagmas, formados a partir de la operación “fusión” (*merge*); en los enfoques basados en el uso,

²¹ Cabe notar que, si bien el ejemplo de Tomasello funciona en inglés, esta construcción transitiva es posible, aunque de manera menos frecuente, en algunas variedades del español. En estas variedades, el uso indica al hablante que es posible generalizar la construcción transitiva al verbo desaparecer y así sucede.

se habla más bien de construcción, que tiene diferentes niveles de abstracción y extensión. En los enfoques basados en el uso, no hay reglas *per se*, sino construcciones que permiten o no la formación de una oración dada. En ambos enfoques, se considera que el léxico se tiene que aprender a partir del *input*, ya que este no forma parte de ningún módulo innato del lenguaje. En el enfoque generativo, el inventario léxico no forma parte del sistema lingüístico central, pero el módulo sintáctico selecciona los elementos del léxico para luego hacer la derivación. En los enfoques basados en el uso, el inventario léxico forma parte de las construcciones; no hay fronteras claras entre las construcciones que representan una sola palabra y una oración completa, por lo que las fronteras entre sintaxis y léxico quedan difuminadas: las oraciones se forman a partir de las construcciones, las cuales pueden estar formadas, a su vez, de otras construcciones.

En las próximas subsecciones, además de exponer las bases teóricas generales sobre los procesos de adquisición de 2L1, L2 y L3, volveremos a comparar las particularidades e implicaciones de dichos procesos según los enfoques generativo y cognitivo.

1.2.2 2L1

Al principio del presente capítulo, presentamos un perfil particular de bi-/multilingüe, el de los hablantes que tienen dos L1, es decir, los que aprenden dos lenguas en su primera infancia (antes de los 3 años). Dos hipótesis han sido postuladas en cuanto a la manera en que se desarrollan las lenguas en el niño bilingüe simultáneo: la de un sistema único (Clark, 2009; Volterra & Taeschner, 1978) y la del desarrollo independiente (De Houwer, 1990; Meisel, 1989, 1994). La primera hipótesis se basa en el principio de contraste, según el cual, en la primera etapa de desarrollo de la lengua, el niño tiene un solo sistema léxico que contiene palabras de las dos lenguas y que no suele contener dos palabras que se refieren a un mismo referente (Volterra & Taeschner, 1978, p. 312). Después de esta primera etapa, las dos lenguas se separan, pero es de esta primera etapa de desarrollo de donde procede la noción de sistema único o unido. Montrul (2013) subraya los problemas metodológicos del estudio que llevó a Volterra y Taeschner a formular su teoría (una base de datos de solo 136 palabras, la orientación del estudio solo hacia el léxico) e insiste en que la ausencia de algo (una palabra en un idioma) en la producción oral de un niño no niega su existencia en su léxico, sino que solo significa que el niño no la usó. Estudios posteriores muestran evidencias a favor de un desarrollo independiente de las dos lenguas, entre otros, los de Barreña

(1997) e Idiázabal (1995) con niños bilingües euskera-castellano (ver también Larrañaga & Guijarro-Fuentes, 2013).

Montrul (2013) no niega que fenómenos como la alternancia de códigos y la interferencia puedan ocurrir en el lenguaje de los niños bilingües, pero se posiciona claramente a favor de la hipótesis del desarrollo independiente. Montrul reitera que el desarrollo del lenguaje en los niños bilingües se hace de manera paralela para cada lengua y, por lo tanto, todos los niños de desarrollo típico, ya sean mono- o bilingües, pasan más o menos por las mismas etapas de desarrollo en cada lengua. Cabe aclarar que esto no significa que los bilingües simultáneos sean comparables en todos aspectos a los monolingües. Por ejemplo, los bilingües simultáneos muestran un leve retraso en la adquisición del léxico, es decir, el niño monolingüe supera al niño bilingüe en el tamaño del léxico cuando se considera solo una de sus lenguas (Bialystok et al., 2010). Sin embargo, esta posible desventaja desaparece cuando se calcula el número de conceptos conocidos –por ejemplo, “sombrero” y “chapeau” representan un mismo concepto– en una o en las dos lenguas del bilingüe e incluso puede darse una ventaja para el bilingüe cuando se suma el número de palabras de las dos lenguas que conoce (De Houwer et al., 2014; Pearson et al., 1993).²²

Si bien hasta aquí hemos centrado la discusión sobre las 2L1 en casos de adquisición en los cuales los dos sistemas lingüísticos se desarrollan paralelamente de manera equilibrada, este no es siempre el caso. De Houwer (2007) llevó a cabo una encuesta en los Países Bajos con cerca de 2000 familias bilingües (neerlandés + una lengua minoritaria) y llegó a la conclusión de que criar niños para que hablen una sola lengua tiene una tasa de éxito de 100 %, excepto en algunos casos de niños de desarrollo atípico. Por su lado, criar niños para que hablen dos lenguas tiene una tasa de éxito del 75 %. Para explicar tales resultados, De Houwer menciona que el *input* de los padres tiene un impacto importante en el desarrollo de las lenguas de un niño criado en un ambiente bilingüe. De todas las diferentes combinaciones de *input* parental, la tasa de éxito más débil aparece cuando uno de los padres solo habla neerlandés y el otro habla a la vez neerlandés y la lengua minoritaria, con una tasa de transmisión de la lengua del 36 %. En cambio, cuando la

²² Por ejemplo, considerando el repertorio léxico de un hipotético bilingüe español-francés en el que cada número corresponde a un concepto único: (1) ‘sombrero’, (2) ‘casa’, (3) ‘hombre’, (4) ‘mujer’, (5) ‘ropa’, (6) ‘comida’, (1) *chapeau*, (2) *maison*, (3) *homme*, (4) *femme*, (7) *jambon* y (8) *gâteau*, el número total de conceptos sería 8 y el número total de palabras conocidas 12 (6 en español, 6 en francés).

situación es totalmente opuesta, es decir, cuando los dos padres hablan la lengua minoritaria y uno habla parcialmente neerlandés, la tasa de transmisión sube a un 93 % y esta tasa sube al 97 % cuando los dos padres solo hablan la lengua minoritaria en casa.

Esto nos trae de vuelta a la noción de *input* y a su papel en la adquisición de una lengua. Recordemos que, en el enfoque generativo, el *input* cumple el papel de desencadenante para que la lengua-I se configure correctamente. No hay consenso sobre la cantidad mínima necesaria para que la adquisición sea exitosa, pero si volvemos al estudio de De Houwer (2007), podemos notar que en el 25 % de las familias que participaron a la encuesta, el *input* proporcionado a los niños en la lengua minoritaria no fue suficiente como para que esa lengua se desarrollara hasta un nivel suficiente. Cabe aclarar que, en el estudio de De Houwer, no se midió el nivel de dominio de los niños en la lengua minoritaria, sino que se preguntó en la encuesta si los niños hablaban o no dicha lengua. En ese sentido, resulta difícil juzgar el nivel de desarrollo real de las lenguas de esos niños. Sin embargo, numerosos estudios se han interesado de manera más detallada en la descripción de las lenguas minoritarias transmitidas en casa por los padres –con mayor o menor éxito– cuando fuera de la casa se habla una lengua mayoritaria. Hasta muy recientemente, se usaba sin cuestionar su validez el término “incompleta” para hablar de la lengua de los hablantes de herencia que no alcanzaban un nivel “nativo” en su lengua de herencia; adquisición incompleta, en oposición a la supuesta adquisición completa en el caso de los nativos.²³ Sin embargo, en los últimos años, Pascual y Cabo y Rothman (2012), Otheguy (2016) y Kupisch y Rothman (2018) han argumentado en contra del uso de este término, lo que provocó una respuesta de Domínguez et al. (2019) y una réplica de Bayram et al. (2019). Domínguez et al. (2019) mantienen que, a pesar de su connotación negativa, el término *incomplete* es irremplazable, ya que caracterizar la gramática de un hablante que no ha adquirido cierto rasgo como “diferente” no es tan preciso como decir que es “incompleta”. Por su parte, Bayram et al. (2019) insisten en el hecho de que la gramática interna de los hablantes de herencia es coherente, a pesar de ser diferente de la norma monolingüe. Los dos principales argumentos en contra del uso del término “incompleto” tienen que ver, por una parte, con la connotación negativa que confiere a las lenguas habladas por esta población ya

²³ Varios investigadores cuestionan la noción de la lengua supuestamente completa que desarrollan los hablantes monolingües (Bayram et al., 2019). Entre los argumentos evocados para poner en duda esta noción, varios autores evocan la heterogeneidad en el nivel de lengua dentro de los monolingües: no todos los hablantes dominan la lengua con la misma destreza.

estigmatizada y, por otra, con el hecho de que no refleja la realidad sobre la naturaleza del sistema lingüístico de estos hablantes. Kupisch y Rothman (2018) proponen “adquisición diferencial” (*differential acquisition*) como alternativa a incompleta, un término que, según sus argumentos, refleja mejor la realidad de los hablantes de herencia. En este sentido, el resultado de la adquisición no es incompleto, sino que difiere de la norma estándar, es decir, de la norma de los hablantes monolingües de la lengua en cuestión. En otras palabras, del mismo modo que los hablantes de una zona pueden tener un dialecto propio de una lengua –que no se considera incompleto sino diferente– los hablantes de herencia desarrollan su propio dialecto divergente de la norma estándar.

Ahora bien, ese debate sobre el término incompleto esconde otra pregunta más importante en el marco de nuestra discusión sobre el papel del *input*: ¿de dónde procede la diferencia entre la lengua de los hablantes de herencia y la de los hablantes monolingües? Una de las fuentes de diferencia procede de la frecuente ausencia de instrucción formal en la lengua de herencia. Kupisch y Rothman (2018) enumeran tres carencias relacionadas con esa fuente de diferencia: (1) hay ciertas nociones que pertenecen a la norma culta y que se enseñan de manera casi exclusiva en la escuela (el subjuntivo en alemán, las reglas de ligación en francés, etc.), (2) el hecho de usar la lengua como medio de instrucción familiariza a los hablantes de esa lengua con los contextos formales en los que se suelen administrar los experimentos lingüísticos y (3) la enseñanza formal expone a los alumnos a una variedad de situaciones de uso de la lengua (literatura, ciencia) y fomenta el desarrollo de habilidades que no se enseñan explícitamente, como la consciencia metalingüística. Si bien la enseñanza formal no es necesaria *per se* para adquirir una lengua, en ausencia de la misma, es necesario proporcionar alternativas equivalentes al niño para que desarrolle su lengua de herencia de manera equilibrada con respecto a la lengua mayoritaria, ya sea mediante la lectura, la exposición a una gran variedad de interlocutores de diferentes registros de habla en diferentes situaciones de la vida, etc.

La segunda fuente de diferencia procede del *input* y del uso de manera más general. Efectivamente, los hablantes de herencia tienen acceso a un *input* diferente del *input* que reciben los monolingües, tanto a nivel cualitativo como cuantitativo: (1) cualitativo porque el *input* es menos diversificado –a menudo solo se habla la lengua en el ambiente de la casa– y a veces contiene muestras de lengua divergentes de la norma monolingüe, ya que el dialecto de los padres se suele ver afectado por el contacto con la lengua mayoritaria y por el fenómeno de atrición (Pascual y Cabo & Rothman, 2012) y (2) cuantitativo porque el *input* que reciben está dividido

entre la lengua mayoritaria y la lengua de herencia, al contrario de los monolingües, que reciben todo su *input* lingüístico en una sola lengua.

La tercera fuente de diferencia tiene que ver con la influencia interlingüística que puede ejercer la lengua mayoritaria sobre la minoritaria (y viceversa). En este caso, las dos lenguas-I de los hablantes de herencia pueden contener innovaciones que no aparecen en el *input* que han recibido.

Este fenómeno se puede explicar desde distintas perspectivas. Desde la lingüística generativa, se considera que, en algunos casos, el *input* proporcionado se aleja lo suficiente de la norma monolingüe como para que la lengua-I del niño se configure de manera divergente. Como prevé la teoría formal, en función del *input* recibido, los parámetros de la GU se van configurando como en el caso de la adquisición monolingüe, pero, en este caso, este proceso se lleva a cabo de manera idiosincrásica ya que el *input* diverge de la norma. Posteriormente, puede ocurrir una influencia interlingüística que va alterando la lengua de herencia. En una visión modular del lenguaje, se considera que los fenómenos ubicados en las interfaces (que tocan más de un componente de la lengua, como la interfaz sintáctico-pragmática) son menos estables porque son más complejos y, por lo tanto, más vulnerables a la influencia interlingüística. Esta propuesta se conoce bajo la hipótesis de la interfaz (*Interface Hypothesis*), término acuñado por Sorace y Filiaci (2006) y ampliamente estudiado en el campo de las lenguas de herencia y del bilingüismo desde una perspectiva generativa (Montrul, 2012; entre otros; Sorace, 2011). En resumen, la lengua-I puede ser diferente desde el principio si el *input* es lo suficiente divergente, o solo diferente a partir del momento en que opera la influencia interlingüística.

Desde la perspectiva de las teorías basadas en el uso, se considera que, dado el *input* reducido, ciertas construcciones no llegan a un nivel de abstracción y atrincheramiento suficientemente desarrollado y son, por lo tanto, menos estables (Zyzik, 2019). Antes de generalizar algunas formas lexicalizadas y hacer abstracciones a partir de ellas, el niño requiere de un *input* suficiente. Zyzik usa el ejemplo del subjuntivo en español, un fenómeno de adquisición tardía, para ilustrar el fenómeno. Un niño que crece en una casa donde se habla español tendrá en su *input* una variedad de oraciones del tipo “cuando salgas ...”, “cuando seas grande...”, “cuando vuelvas...”, pero al principio solo aprenderá esas construcciones de manera fija, lexicalizada. Es decir que, a pesar de dominar esas construcciones lexicalizadas y poder usarlas en los contextos

adecuados, en un primer momento, no será capaz de generalizarlas a otros verbos y no tendrá la construcción más abstracta (Adv + V_{Subjuntivo}). Si a lo largo de su proceso de adquisición el niño no tiene acceso a un *input* que contenga una cantidad suficiente de muestras de esa estructura, esta no llegará a ser lo suficientemente estable y se podrán observar variaciones en la producción del niño.

Hasta aquí, a través de los ejemplos del repertorio léxico y de los hablantes de herencia, hemos podido constatar que “un bilingüe no equivale a dos monolingües en una misma persona” (Grosjean, 1989, p. 3). Al margen de que el proceso de adquisición de las dos lenguas se haga de forma relativamente independiente, el bilingüe maneja dos sistemas permanentemente (Starreveld et al., 2014) y tiene que inhibir una lengua para usar la otra, lo que conlleva un cierto coste cognitivo. La lista de diferencias cognitivas entre monolingües y bilingües es larga y no es nuestro objetivo presentarla de manera exhaustiva (véase, por ejemplo, Bialystok et al., 2012). Sin embargo, a modo de conclusión provisional, cabe subrayar que esas diferencias ponen en evidencia la necesidad de usar criterios específicos para evaluar, de manera general, las competencias de los bilingües, por un lado, y de los monolingües, por otro.

1.2.3 L2

Como hemos mencionado anteriormente, nos referimos a la adquisición de L2 cuando el hablante empieza la adquisición una lengua adicional después de los tres años. A su vez, podemos separar a los hablantes de L2 en tempranos o tardíos según si han adquirido su L2 antes o después de la pubertad (Montrul, 2013). Vemos, por lo tanto, que el factor edad es uno de los aspectos más relevantes en lo que respecta a la AL2. Existe la creencia de que los hablantes que adquieren su L2 más jóvenes –adquisición temprana– tienen más éxito, mientras que los casos de adultos que fracasan o que no llegan a un nivel satisfactorio a la hora de adquirir una L2 son más numerosos. Estas creencias se basan en la hipótesis del periodo crítico (Lenneberg, 1967) que, en principio, se formuló para la adquisición de la L1. Dicha hipótesis sostiene que es imposible adquirir el lenguaje después de la pubertad. Si bien hay cierto consenso al respecto de la existencia de un periodo crítico en el caso de la AL1, su impacto en lo que respecta a la AL2 es más controvertido (Roca de Larios & Manchón Ruiz, 2006, p. 64). Existen, por ejemplo, casos de adultos que han adquirido su L2 de manera tardía y que alcanzan niveles muy avanzados de competencia (véase, por ejemplo, Bongaerts, 1999).

En su repaso de los estudios sobre el factor edad en la adquisición de la L2, Roca de Larios y Manchón Ruiz (2006) encuentran pocas evidencias a favor de la supuesta superioridad de los niños para adquirir una L2, por lo menos en un marco de educación formal. Efectivamente, varios estudios presentados por los autores (Burstall et al., 1974; Cenoz, 2003b; García Lecumberri & Gallardo, 2003; García Mayo, 2003) indican que, con el mismo tiempo de enseñanza en un contexto formal, los aprendientes que habían empezado su adquisición más tarde tenían mejores resultados que los que habían empezado antes. Si bien los autores no niegan que los niños puedan aprender mejor una lengua en un contexto natural, también afirman que no se puede generalizar esta facilidad a todos los contextos y que, por lo tanto, los niños no son, en sí, mejores aprendientes de L2 que los adultos. Al respecto, Hoffman (2014) afirma que “no es posible encontrar pruebas sólidas de que los niños sean mejores aprendientes que los adultos” ya que “no se pueden comparar cosas incomparables” (p. 35). Con “cosas incomparables”, Hoffman se refiere a las grandes diferencias que existen entre un niño y un adulto en varios aspectos, como el desarrollo cognitivo, su capacidad analítica y su experiencia en el aprendizaje.

En cualquier caso, al aprender su L2, los hablantes desarrollan un nuevo sistema lingüístico que se conoce como “interlengua” (*interlanguage*), término propuesto por Selinker (1972) para referirse al sistema lingüístico en desarrollo en la mente del aprendiente de L2 (o de manera general, de cualquier LNN). La formación de la interlengua se basa en cinco procesos cognitivos principales: las transferencias a partir de la L1, las transferencias relacionadas con el aprendizaje, las estrategias de aprendizaje, las estrategias de comunicación y la sobregeneralización de las reglas de la lengua meta. La interlengua se caracteriza por ser idiosincrásica (propia de cada aprendiente) y sistemática (no varía de manera aleatoria, sino que se rige por sus propias reglas). A lo largo de los años, la hipótesis de la interlengua ha generado varios estudios en el campo de la AL2 (véase, a modo recapitulativo, Han & Tarone, 2014) y ha suscitado debates sobre cuestiones como la fosilización de los errores, entre otros. Efectivamente, la fosilización –entendida como la estabilización permanente no deseada de ciertos rasgos de la interlengua que no corresponden a la lengua meta²⁴– ha sido examinada de manera detallada para encontrar las posibles causas y

²⁴ La estabilización representa un paso hacia la fosilización. La simple estabilización se puede desbloquear para que el aprendiente alcance la forma de la lengua meta, pero a partir del momento en que deja de haber variación entre la forma correcta e incorrecta, se dice que hay fosilización. La fosilización es, entonces, irreversible (Han, 2013, p. 140).

características de dicho proceso (Selinker & Lamendella, 1978). Investigaciones más recientes como las de Han (2004, 2013) permitieron visitar la hipótesis inicial para formular la hipótesis de la fosilización selectiva, que postula cuatro generalizaciones: (1) la fosilización es selectiva, (2) la fosilización afecta a la adquisición de una variedad de estructuras de la lengua meta relacionadas con la forma, el significado y las funciones (pragmático-discursivas), (3) la fosilización está condicionada por la manera de pensar en la L1, pero está influenciada y reforzada por las características del *input* en la L2 y (4) la fosilización es más observable en la producción espontánea, en la que el aprendiente tiene que formular sus ideas y expresarse a su manera (Han, 2013, p. 165). A pesar de las puntualizaciones aportadas a la hipótesis inicial, la noción de interlengua de Selinker sigue siendo muy útil para referirse al estado transitorio de la L2 o las LNN de manera general.

En el enfoque generativo, como sucede en el campo de la adquisición de la L1, son numerosas las investigaciones sobre adquisición de la L2. Uno de los grandes debates sobre el proceso de adquisición de L2 dentro del marco del generativismo es la posibilidad de acceso a la GU durante el proceso de adquisición de una L2 (White, 2003, p. 15). Por un lado, existen hipótesis que postulan que este acceso a la GU es posible (como sucede en la adquisición de L1) y, por otro, tenemos hipótesis que defienden la idea contraria, es decir, la imposibilidad de acceder directamente a la GU en el proceso de adquisición de una L2. Una de las primeras hipótesis que se adscriben a este segundo punto de vista es la llamada hipótesis de la diferencia fundamental (*Fundamental Difference Hypothesis*; Bley-Vroman, 1989). Uno de los argumentos clave de esta hipótesis procede de la diferencia en cuanto al nivel alcanzado en L1 y en L2. Si bien se puede decir que, en un contexto de desarrollo típico, cualquier hablante culmina el proceso de adquisición de su L1 con éxito, pocos son los aprendientes adultos de L2 que llegan a un nivel casi nativo en su L2 (Bley-Vroman, 1989, p. 44). Este contraste lo atribuye Bley-Vroman a la posibilidad de acceder directamente a la GU en el caso de la L1, lo cual garantiza el éxito del proceso de adquisición, mientras que, para la AL2, tal acceso directo ya no existe. La hipótesis de la diferencia fundamental propone que, para el aprendiente adulto de L2, la GU solo es accesible a través de las configuraciones de la L1. De manera similar, la hipótesis de los rasgos funcionales fallidos (*Failed Functional Features Hypothesis*; Hawkins & Chan, 1997) predice que, después de la pubertad, resulta imposible adquirir nuevos rasgos abstractos, es decir, rasgos abstractos que no han sido instanciados en la L1.

Todas las demás hipótesis que presentamos a continuación consideran que la GU sigue siendo accesible de una manera u otra. Dependiendo de lo que consideran como el estado inicial de la L2, estas hipótesis se dividen en dos categorías que podríamos denominar: (1) las de la transferencia y (2) las de la no transferencia.

En el primer grupo se encuentra la hipótesis de la transferencia completa y del acceso completo (*Full Transfer Full Access Hypothesis*; Schwartz & Sprouse, 1996). Esta hipótesis predice que el estado inicial de la L2 es una copia de la L1 (transferencia completa) y que, a medida que el proceso de adquisición avanza y el aprendiente va recibiendo *input* en su L2, este puede acceder a la GU (acceso completo) para corregir los errores de configuración del estado inicial de la L2. Por su lado, la hipótesis de los árboles mínimos (*Minimal Trees Hypothesis*; Vainikka & Young-Scholten, 1996b, 1996a) no predice una transferencia completa, sino únicamente de las categorías léxicas. Las categorías funcionales no se transfieren a partir de la L1 (ni en el estado inicial, ni después), sino que se adquieren a partir del *input* que el aprendiente recibe en su L2. Sin categorías funcionales, al principio, el aprendiente no puede proyectar árboles sintácticos más allá de las categorías léxicas, de ahí el nombre de “árboles mínimos”. Esta hipótesis supone que, en cuanto a las categorías funcionales, no hay diferencia entre la AL1 y la AL2, es decir que se desarrollan de manera gradual. Otra propuesta es la de la hipótesis de los rasgos sin valores definidos (*Valueless Features Hypothesis*; Eubank, 1993, 1996), que predice una transferencia tanto de las categorías léxicas como funcionales pero que, en el caso de las categorías funcionales, la transferencia de los rasgos se realiza sin valores definidos, lo que permite cierta opcionalidad en la interlengua de los aprendientes de L2.²⁵

En el segundo grupo se encuentran las hipótesis que consideran que el estado inicial de la L2 no es la L1, sino una gramática virgen de cualquier huella de la L1. La hipótesis inicial de la sintaxis (*Initial Hypothesis of Syntax*; Platzack, 1996) y la hipótesis del acceso completo sin transferencia (*Full Access Hypothesis*; Epstein et al., 1996) son dos hipótesis que presentan la no transferencia como el estado inicial de la L2.

²⁵ Como subraya White (2003, pp. 86-87), en la hipótesis de Eubank (1993, 1996), la noción misma de rasgos sin valores definidos no queda del todo clara. Además, White no encuentra motivos teóricos para postular este estado ‘inerte’ de los rasgos.

A su vez, la hipótesis de la estructura superficial (*Shallow Structure Hypothesis*; Clahsen & Felser, 2006a, 2006b) y la hipótesis de la flexión superficial ausente (*Missing Surface Inflection Hypothesis*; Haznedar & Schwartz, 1997; Prévost & White, 2000) no se enfocan en el acceso a la GU de manera directa, sino en las diferencias de procesamiento en la L2. Según la hipótesis de la flexión superficial ausente, la interlengua del hablante de L2 puede tener una representación adecuada de los rasgos funcionales abstractos necesarios para procesar la sintaxis, pero puede presentar problemas para establecer la relación entre el rasgo abstracto y la forma, lo que resulta en la agramaticalidad al nivel superficial. Según esta hipótesis, estos problemas de relación entre rasgo y forma suelen aparecer de manera más obvia en la L2 en las tareas de producción oral, ya que son más exigentes a nivel cognitivo. Con un resultado similar, la hipótesis de la estructura superficial postula que existen dos vías para interpretar y producir información en una lengua: el análisis sintáctico completo (*full parsing*, que depende del sistema gramatical) y el procesamiento superficial (*shallow processing*, que depende del léxico, de la semántica, de la pragmática, del conocimiento general y de los mecanismos de asociación significado-forma). En la L1, se pueden usar las dos vías: el procesamiento superficial permite tratar la información de manera más rápida y el análisis sintáctico permite confirmar la información en caso de ambigüedad. En cambio, en la L2, predomina la vía del procesamiento superficial, por lo que se pueden observar diferencias en el procesamiento, particularmente en las estructuras complejas. Esta hipótesis predice que, si bien el hablante de L2 puede alcanzar un nivel prácticamente comparable al de un nativo, deberían persistir ligeras diferencias en el procesamiento.

Hasta aquí hemos presentado hipótesis que se adscriben a la perspectiva generativa, pero en las perspectivas basadas en el uso también encontramos trabajos sobre la adquisición de segundas lenguas. Según las perspectivas basadas en el uso, las diferencias en cuanto al nivel alcanzado en la L2 en comparación con la L1 se atribuyen al contexto y a la experiencia de adquisición. Mientras que el niño asimila una lengua, unos gestos, unas posturas y unas expresiones faciales de su cultura al mínimo detalle, el aprendiente de L2 es un ser humano que ya sabe comunicarse en su L1 y que intenta replicar esto en una L2 (Bybee, 2008, p. 232). El hecho de tener conocimientos lingüísticos anteriores atrincherados de manera firme en la mente condiciona de cierta manera la adquisición de una L2, ya que la L1 proporciona un punto de partida para comparar las nuevas construcciones de la L2. Según Bybee, en la medida en que las construcciones son idénticas entre la L1 y la L2, las perspectivas basadas en el uso predicen un

papel facilitador de la L1. En cambio, si hay diferencias, por mínimas que sean, estas representarán un obstáculo para la adquisición de dichas construcciones. Sin embargo, con un *input* y una práctica suficientes, todas las construcciones de la L2 pueden llegar a ser dominadas por el aprendiente de L2. Del mismo modo que se produce la adquisición de la L1 –construcción por construcción– la L2 se desarrolla paulatinamente a medida que el aprendiente va categorizando las construcciones que recibe en su *input* (Ellis & Cadierno, 2009; Ellis & Collins, 2009), un desarrollo estrechamente vinculado con la frecuencia (Ellis, 2002), la fiabilidad de los indicios (MacWhinney, 1997) y la prominencia de esos indicios (Ellis, 2006). Dado que los aprendientes de L2 son normalmente más maduros cognitivamente, los contextos y las condiciones de aprendizaje suelen ser muy diferentes en comparación con un niño que adquiere su L1.

La gran mayoría de las propuestas de adquisición de L2 que hemos presentado consideran que la L1 tiene un impacto en la adquisición de la L2. Como veremos en la sección 1.4, los modelos e hipótesis de adquisición de L3 presuponen también cierta influencia a partir de la L1 o L2 (o ambas), ya que se inspiran en gran medida de las propuestas de AL2. En la línea de las propuestas sobre AL3 que se presentarán más adelante, este trabajo presupone también, *a priori*, la existencia de algún tipo de influencia interlingüística, al menos en las etapas iniciales de adquisición de una LNN.

1.3 Adquisición de L3/Ln

1.3.1 L3

Como hemos visto en la sección anterior, en los estudios del campo de AL2, el contraste se establece entre la L1 y la lengua objeto de estudio, que se presenta bajo la etiqueta de L2. Sin embargo, como subraya De Angelis (2007), en estos estudios, la frecuentemente denominada L2 no es necesariamente la segunda lengua de los participantes ni en orden de adquisición, ni en orden de dominio, sino, a veces, la tercera, cuarta, etc.. Por ejemplo, a menudo el español no es la “segunda” lengua de los aprendientes que provienen de Quebec, ya que, de manera general, tienen que aprender francés e inglés antes de tomar cursos de español. Esta observación se aplica a varios países o regiones multilingües donde se han llevado a cabo estudios de AL2 (por ejemplo, Canadá, Suiza, Italia, Bélgica, Cataluña, el País Vasco, etc.). Es más, considerando la presencia del inglés como lengua de comunicación internacional, esta observación se aplica también, de manera más

global, a los hablantes de zonas no anglófonas con un cierto nivel de escolarización que tienen el inglés como L2N. En todos estos casos, que son cada vez más numerosos, consideramos pertinente la necesidad de hacer una distinción entre la adquisición de una L3 y la de una L2.

Para nuestro estudio, adoptamos la definición de Cenoz (2003a), que presenta la adquisición de una L3 como la adquisición de una lengua no-nativa por aprendientes que han adquirido previamente, o que están adquiriendo, otras dos lenguas. La adquisición de las otras dos lenguas puede ser simultánea (como en el bilingüismo temprano) o consecutiva (p. 71). Esta definición implica que, si un hablante tiene dos lenguas adquiridas de manera simultánea en la infancia temprana y otra más tarde, no tiene ninguna L2, sino dos L1 y una L3. Tal visión contrasta con las ideas expuestas en Hammarberg (2010), donde se considera que una lengua adquirida después de dos lenguas adquiridas simultáneamente en la infancia (2L1) no es un verdadero caso de L3, sino un caso de L2. A continuación, en la Figura 1.2, reproducimos la jerarquía presentada por Hammarberg para delimitar lo que son las lenguas primarias, secundarias y terciarias.

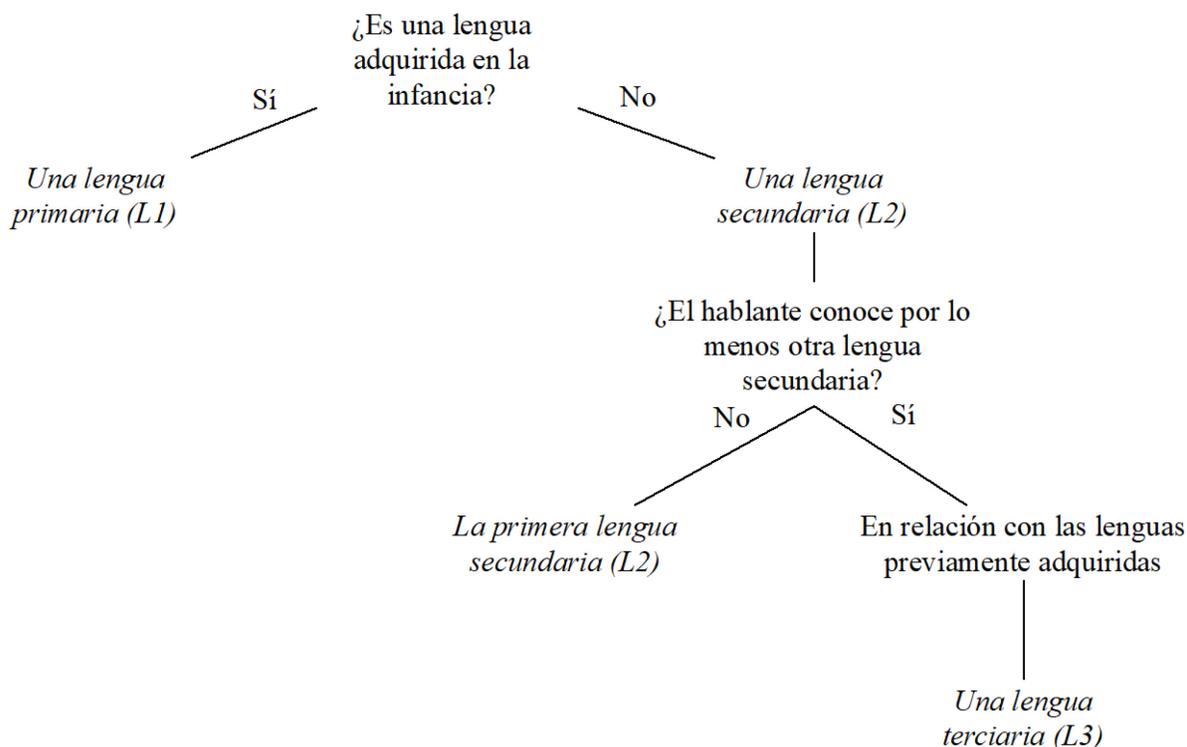


Figura 1.2. Jerarquía para clasificar las lenguas adquiridas, traducida de Hammarberg (2010, p. 101)

Hammarberg aboga por el uso de las expresiones *primaria*, *secundaria* y *terciaria* para romper con el orden cronológico evocado por los términos *primera*, *segunda* y *tercera* e insistir más bien en el papel cognitivo que cumple cada categoría de lengua. Ese leve cambio de nomenclatura permite conservar las mismas etiquetas L1, L2, L3. Si bien Hammarberg considera que todas las lenguas adquiridas después de la infancia se agrupan bajo la etiqueta L2, también sostiene que dentro de este grupo hay una subcategorización. Dicho de otro modo, todas las L3 forman parte del gran grupo de L2, pero no todas las L2 son a la vez L3. El autor también hace referencia a la problemática de clasificar a los hablantes que conocen más de tres lenguas. En estos casos, la lengua que es objeto del estudio –en relación con las demás– se debe considerar como la L3 y todas las lenguas no nativas, adquiridas antes o después, se deben considerar como L2. El modelo de Hammarberg tiene sus ventajas, como el hecho de limitar los niveles a tres sin tener que hablar de L4, L5, L6, pero tiene una limitación importante en cuanto a los bilingües simultáneos. Según su modelo, la tercera lengua adquirida (cronológicamente) por un bilingüe simultáneo no representa ninguna diferencia en comparación con la segunda lengua de un bilingüe secuencial: ambas son consideradas como L2. Esta limitación sobresale cuando consideramos una diferencia importante entre los aprendientes de L3 (según Cenoz) y los aprendientes de L2 (que solo tienen una L1). Los aprendientes de L3 tienen acceso, de manera consciente o no, a dos lenguas en vez de una y tienen mayor experiencia en la adquisición de lenguas, ya se haya realizado esta adquisición en un contexto natural (en la infancia temprana) o más formal (después de los 3 años). Reiteramos entonces nuestra preferencia por la definición de Cenoz (2003a) en el marco del presente trabajo. Esta definición se aplica de manera adecuada a las regiones que tienen una lengua minoritaria, donde varios de los hablantes de la lengua minoritaria son bilingües simultáneos, como es el caso de nuestro estudio.

Volviendo al hecho de conocer dos lenguas a la hora de adquirir la tercera, lógicamente, se podría valorar este acceso a dos lenguas tanto de manera positiva (experiencias de aprendizaje más diversas, mayor repertorio léxico, sintáctico, fonológico, morfológico, etc.) como negativa (posibilidades de interferencia de dos lenguas en vez de una). Sin embargo, varios estudios (ver detalles en Bowden et al., 2005) defienden la tesis de que los multilingües –entre ellos, los aprendientes de L3– aprenden mejor y más rápido que los monolingües. Es decir que, si en algunos casos se notan efectos negativos (interferencia) de las lenguas anteriores, los beneficios aportados superan dichos efectos.

Ahora bien, del mismo modo que existen diferencias entre la adquisición de L2 y de L3, podríamos argumentar que existen diferencias similares cuando se compara la L3 a una posible L4, L5 o L_n (n denota cualquier lengua adicional después de la L3). Sin embargo, consideramos, como menciona Hufeisen (1998), que la diferencia entre una L3 y una L_n es más cuantitativa que cualitativa, es decir que, si bien aumentan las fuentes posibles de influencia, solo se añaden otras LNN más en la ecuación sin que cambien las variables. Es posible que, en los próximos años, nuevos descubrimientos sobre los procesos de adquisición de lenguas adicionales nos empujen a adoptar una nueva terminología para distinguir, por ejemplo, la L4 de la L3, pero por el momento, con el estado actual de la ciencia, concluiremos que la etiqueta L3 es válida para abarcar todas las lenguas adquiridas después de dos lenguas. A continuación, exploramos algunos constructos que ayudan a definir las particularidades de los aprendientes de L3.

1.3.2 Nociones clave de AL3

Como subrayan Rothman (2010) y García-Mayo (2012), los datos procedentes de estudios de AL3 permiten revisar cuestiones de AL2 desde una nueva perspectiva y es precisamente lo que ocurre con la influencia interlingüística. Desde los inicios de la AL3 como campo de estudio, la noción de influencia interlingüística ha ocupado un lugar central, ya que las fuentes posibles de influencia hacia la lengua meta son múltiples y no solo se restringen a la L1, como en el caso de la adquisición de L2.²⁶ Por lo tanto, no es extraño encontrar que el objetivo principal de la mayoría de los estudios de AL3 consista precisamente en determinar la fuente de dicha influencia, ya sea esta la L1, la L2 o ambas. Una de las primeras áreas de interés en AL3 fue el léxico, como muestra el estudio pionero de Ringbom (1987) que versa esencialmente sobre las innovaciones léxicas y los préstamos en inglés como L3 por parte de hablantes de finés (L1 o L2) y sueco (L1 o L2). El volumen editado por Cenoz, Hufeisen y Jessner (2001) representa otro ejemplo donde destaca el interés por el léxico, con una mayoría de capítulos que versan sobre distintos aspectos del acceso y del procesamiento del léxico (la transferencia léxica hacia la L3, el fenómeno de la “punta de la lengua” en la L3 y la organización del lexicón plurilingüe, entre otros). Poco después, surgió el interés por la influencia en la morfosintaxis con trabajos sobre varios fenómenos como las

²⁶ El presente apartado se limitará a la influencia interlingüística en AL3, invitamos al lector a volver a la sección 1.1.5 del presente capítulo para una definición más general del concepto.

cláusulas relativas en L3 inglés (Flynn et al., 2004), el uso de rellenos (*placeholders*) en L3 inglés (Mayo et al., 2005), el sintagma determinante en L3 francés (Leung, 2005), la concordancia con los pronombres en L3 inglés (García Mayo et al., 2006) o la negación en L3 neerlandés/sueco (Bardel & Falk, 2007). La transferencia morfosintáctica sigue llamando la atención de varios investigadores de tal modo que, en una revisión exhaustiva de la literatura sobre el tema, Puig-Mayenco, Gonzalez Alonso y Rothman (2020) reportan nada menos que 71 estudios publicados entre 2004 y 2017 sobre los factores de transferencia en la adquisición de la morfosintaxis en una L3. El interés por la influencia a nivel fonológico y fonético despertó más tarde, como menciona Cabrelli Amaro (2012) en un artículo sobre el estado de la cuestión. Cabrelli Amaro nota que el tema ganó cierta importancia con la publicación de un número especial de la revista *International Journal of Multilingualism* sobre este tema (febrero 2010).

A continuación, exploramos los factores más estudiados que suelen condicionar la influencia interlingüística.

1.3.2.1 Edad, contexto y orden de adquisición

No todos los aprendientes de L3 son iguales y sus perfiles varían, entre otras cosas, según el contexto y la edad de adquisición de sus lenguas previas. Por ejemplo, el mero hecho de encontrar invertida la combinación de las dos primeras lenguas en dos hablantes distintos cambia de manera considerable su perfil como aprendientes de L3, aunque tengan un nivel muy avanzado en las dos lenguas previas. Como vimos al principio del presente capítulo, cuanto más temprano en la vida del niño se adquiera la L2, más probable es que la adquisición se dé en un contexto natural y que se trate, por lo tanto, de una adquisición vinculada a conocimientos más implícitos y automatizados. En contraste, una lengua adquirida después de los 3 años –una L2– tiene más probabilidades de ser adquirida en un contexto formal y de estar vinculada a conocimientos más explícitos y menos automatizados. Considerando esto y todo lo que hemos visto hasta aquí en el presente capítulo, las representaciones de cada lengua en la mente del hablante no serán iguales según se trate de una L1/2L1 o de una L2. Ciertos estudios, como los de Carvalho y da Silva (2006) y Koike y Gualda (2008), muestran que, a la hora de adquirir una L3 en un contexto formal, el contexto de adquisición de las lenguas previas tiene un impacto importante. En estos estudios sobre la adquisición del portugués como L3, los aprendientes que habían adquirido el español L2 en un contexto formal mostraron mayor sensibilidad a los análisis contrastivos y la enseñanza explícita

que los que tenían 2L1 (inglés-español) y los que tenían el español como L1. Además, un estudio de Johnson (2004), también sobre aprendientes de portugués, como L3 mostró que los aprendientes que tenían el español como L1 o como lengua de herencia cometían ciertos errores ortográficos que los que tenían el español como L2 no cometían. Estos estudios sugieren que el tipo de conocimiento desarrollado en las L1/2L1/L2 cambia de manera significativa la manera de adquirir y usar la L3. A continuación, exploramos más en detalle un constructo vinculado con los conocimientos de las lenguas, el de conocimiento metalingüístico.

1.3.2.2 Conocimiento metalingüístico

Por definición, todos los hablantes tienen conocimientos lingüísticos que les permiten comunicarse en una lengua dada. Luego, todos tienen, hasta cierto punto, conocimientos sobre la(s) lengua(s) que hablan, unos conocimientos que no son necesarios en sí para comunicarse, sino para reflexionar sobre la lengua; esos son los llamados conocimientos metalingüísticos. En la literatura sobre el tema, más allá del término “conocimiento” se usan también términos como “destrezas” (*abilities/skills*) y “consciencia” (*awareness*) metalingüísticas para referirse a la puesta en práctica de dichos conocimientos. Así, Bialistok y Barac (2013) definen las destrezas metalingüísticas como el conocimiento explícito de las estructuras lingüísticas y la habilidad de acceder a este conocimiento (p. 193). De manera similar, Pratt y Grieve (1984, p. 2), definen la consciencia metalingüística como la habilidad de reflexionar sobre la naturaleza y las funciones del lenguaje.

Se considera que los bi-/multilingües, por sus ricas y diversas experiencias lingüísticas tienen mayores conocimientos metalingüísticos, y, por lo tanto, una consciencia metalingüística más desarrollada. Varios autores, como Thomas (1988), asocian el desarrollo más avanzado de la consciencia metalingüística a una mayor capacidad de adquisición de las lenguas. Según Thomas, se trata de una herramienta que puede llevar al aprendiente a comprender de manera más rápida y eficaz los nuevos elementos de su L3 y, de esta manera, acelerar su proceso de adquisición. En el mismo sentido, Herdina y Jessner (2002) hablan del factor M(ultilingüismo) para diferenciar al hablante multilingüe del monolingüe. Bajo el factor M, los autores agrupan diferentes características, pero una de las más importantes es la consciencia metalingüística, entendida como un monitor en un sentido similar a la idea de Krashen (1981), es decir, una herramienta que ayuda al aprendiente a procesar y producir de manera más óptima en sus diversas lenguas. Si bien

Krashen habla del monitor como una herramienta disponible a la hora de adquirir una L2, Herdina y Jessner se refieren a este concepto bajo el nombre de monitor mejorado del multilingüe (*Enhanced Multilingual Monitor*), que aumenta, entre otras cosas, la habilidad de aprender, controlar y mantener las lenguas.

Existe cierto consenso en cuanto al efecto beneficioso de la consciencia metalingüística para la adquisición de L3, pero, a menudo, esta pasa desapercibida en la formulación de los principales modelos de influencia interlingüística en adquisición de L3 porque es difícil de definir y de cuantificar. Ante la dificultad de definirla de una manera suficientemente objetiva, resulta difícil operativizarla en un modelo y así establecer predicciones de influencia interlingüística en adquisición de L3. El modelo dinámico del multilingüismo (*Dynamic Model of Multilingualism*; DMM; Herdina & Jessner, 2002) representa una excepción notable en el panorama de los modelos de adquisición de L3, ya que ubica el factor M en el centro del modelo. Sin embargo, esta propuesta se queda en un nivel muy abstracto al predecir una fluctuación de las lenguas de un multilingüe de manera no-lineal y compleja, sin ofrecer predicciones claras sobre los posibles casos concretos de influencia de la L1 y la L2 sobre la L3. Por lo tanto, no incluiremos este modelo en nuestra revisión de los modelos de influencia en AL3 en la sección 1.4.

1.3.2.3 Distancia tipológica y psicotipológica

La distancia entre las lenguas es un factor bien estudiado y, a pesar de las diversas maneras de concebir dicha distancia, resulta relativamente fácil operacionalizar esta variable en un modelo. Podemos separar las diferentes maneras de considerar la distancia en dos grandes categorías. Por un lado, la tipología lingüística consiste en la clasificación de las lenguas según ciertos criterios objetivos generalmente establecidos por los lingüistas y, por otro, la psicotipología, tal como la define Kellerman (1983), consiste en la clasificación de las lenguas según las percibe el aprendiente. A continuación, nos centramos en la noción de tipología para luego interesarnos en la de psicotipología.

De manera general, los criterios utilizados para establecer la distancia tipológica son de tres tipos: genéticos, formales y geográficos. En la mayoría de los estudios, proximidad tipológica tiende a ser equiparada a proximidad genética, es decir, dentro de las familias de lenguas. Por ejemplo, Liceras y Alba de la Fuente (2015) consideran que, cuando se habla de *proximidad* tipológica, se deben considerar los vínculos genéticos entre las lenguas, ya que estos son

responsables de la variación conforme a los macro-parámetros de las lenguas, como el orden canónico de los componentes o el tipo de morfología. En cambio, la *semejanza* tipológica se establece en función de las relaciones formales entre las lenguas, por ejemplo, el parámetro del sujeto nulo, el orden de los elementos de un grupo clítico o la posición pre- o posverbal del clítico complemento de un verbo en infinitivo.²⁷ Según estos criterios, el francés y el español serían lenguas tipológicamente cercanas, pero no semejantes en cuanto a ciertas estructuras. Los ejemplos (1.6) a (1.8), tomados de Licerias y Alba de la Fuente (2015, pp. 335-336), muestran como las dos lenguas difieren en cuanto a la posición de los clíticos:

- (1.6) a. ¿Los libros? Se[DAT] los[ACC] enviaré mañana
- b. *Les livres? Je lui[DAT] les[DAT] enverrai demain
- c. Les livres? Je les[ACC] lui[DAT] enverrai demain
- (1.7) a. Quiero verla
- b. *Je veux voir la
- c. Je veux la voir
- (1.8) a. Yo la quiero ver
- b. *Je la veux voir
- c. Je veux la voir

Entonces, a pesar de tener una proximidad tipológica (de manera global), en este caso, el francés y el español no siempre presentan semejanzas tipológicas (de manera más específica). Del mismo modo, dos lenguas que no tienen proximidad tipológica podrían tener semejanza tipológica en cuanto a ciertos aspectos de la lengua. Por ejemplo, el euskera y el español son dos lenguas tipológicamente alejadas (el euskera es una lengua aislada, sin vínculo genético con el español), pero presentan semejanzas tipológicas cuando las comparamos a nivel estructural: ambas lenguas tienen el valor [+sujeto nulo]. Según Comrie (1989), las semejanzas tipológicas se pueden atribuir a cuatro razones: el azar, los vínculos genéticos, los universales lingüísticos o la proximidad geográfica. La última razón evocada por Comrie nos permite volver sobre la distancia tipológica vista desde el criterio geográfico, un criterio no muy estudiado en los trabajos de AL3, pero que sí tiene cierta importancia en los estudios que versan sobre la evolución diacrónica de las lenguas y el contacto de lenguas. Concretamente, por cuestiones de proximidad geográfica y de contacto entre las lenguas, dos lenguas genéticamente no relacionadas y formalmente muy distintas podrían llegar a compartir ciertas características formales. Si bien, a priori, este criterio geográfico no

²⁷ Esta noción de semejanza tipológica equivale a lo que Westergaard et al. llaman semejanza estructural (2017), que exploraremos de manera más detallada al hablar de su modelo, el *Linguistic Proximity Model*, en la sección 1.4.5.

parece relevante para los estudios de AL3, puede tener cierto impacto en cuanto a la percepción de la proximidad de las lenguas por el hablante.

Con este concepto de proximidad percibida por el hablante, volvemos a la noción de proximidad psicotipológica introducida por Kellerman (1983). La distancia psicotipológica es subjetiva y puede diferir hasta cierto punto de un aprendiente a otro. Por ejemplo, si un hablante tiene que establecer el grado de proximidad entre tres lenguas como el francés, el japonés y el español, es muy probable que encuentre más proximidad entre el francés y el español, dos lenguas romances. En cambio, si uno tiene que juzgar la distancia entre el francés, el italiano y el español, tres lenguas romances, resulta más difícil predecir cuáles de las tres lenguas parecerán más cercanas para el hablante.

Varios de los modelos de adquisición que presentaremos en la sección 1.4 consideran el concepto de distancia entre las lenguas, ya sea a nivel estructural o tipológico, como factor determinante en el proceso de adquisición de L3.

1.3.2.4 Asociación de extranjería (*foreignness*) o el estatus de la(s) L2

De Angelis (2005) afirma que los aprendientes multilingües asignan (de manera inconsciente) el estatus de “lengua extranjera” a sus LNN, lo que resulta en un vínculo cognitivo más fuerte entre las LNN que entre una LN y una LNN. Así, para no sonar como si estuviera hablando su LN, el hablante multilingüe evita recurrir a su LN y favorece la influencia de otra LNN a la hora de usar su L3. La asociación de extranjería resulta entonces en una mayor aceptación de la influencia a partir de una LNN que de la(s) LN.

En la misma línea, Williams y Hammarberg (1998) añaden que las LNN son más susceptibles de estar seleccionadas como fuentes de influencia en las etapas iniciales de adquisición de una L3 por cuestiones relacionadas a las diferencias entre el proceso de adquisición de la L1 en oposición a la adquisición de una L2 (que, como lo hemos visto al principio del capítulo se adquiere en un contexto más formal). Dadas estas diferencias, a la hora de adquirir una L3, se reactivan los mismos mecanismos que se usaron para la adquisición de la L2 y, por la tanto, la influencia a partir de la LNN es más susceptible de ocurrir.

1.4 Modelos de adquisición de L3

A partir de los factores de influencia interlingüística presentados en la sección 1.3.2, varias hipótesis y modelos han sido propuestos en las dos últimas décadas. En la presente sección, veremos las diferentes posibilidades de transferencia y los modelos que han recibido más atención para cada posibilidad. Empezaremos con las dos propuestas que otorgan mayor importancia al orden de adquisición de las L1 y L2, es decir, la transferencia exclusiva a partir de la L1 o de la L2. Luego, continuaremos con propuestas que consideran la proximidad entre las lenguas como factor más importante y veremos cómo dicha proximidad puede ser concebida en varios niveles en cada modelo presentado. Todos los modelos mencionados hasta ahora se conforman al marco generativo, pero el campo de AL3 también se ha estudiado desde las aproximaciones basadas en el uso, por lo que, para terminar la revisión de los modelos, presentaremos algunas ideas de propuestas basadas en el uso, así como algunos factores que se consideran importantes desde esta perspectiva.

1.4.1 Hipótesis del factor de la L1

Como hemos visto en la sección 1.2.3, en el caso de la adquisición de la L2, la única transferencia posible procede de la L1 (de aquí en adelante, *L1 factor*).²⁸ En el caso de la adquisición de una L3, la transferencia exclusiva a partir de la L1 también ha sido considerada como una opción posible. Na Ranong y Leung (2009) estudiaron la adquisición de los objetos nulos en chino mandarín por parte de hablantes de L1 tailandés/L2 inglés (L3 chino mandarín) y también de un grupo de L1 inglés (L2 chino mandarín). La distribución de los objetos nulos y explícitos en tailandés y en chino mandarín es comparable, pero diferente del inglés. Por lo tanto, el diseño del estudio se prestaba a tres posibilidades de transferencia: (1) facilitadora a partir de la L1 tailandés o (2) no facilitadora a partir de la L2 inglés en el caso de los aprendientes de L3 chino mandarín y (3) no facilitadora a partir de la L1 inglés en el caso del grupo de L2 chino mandarín. Las autoras observaron una mayor sensibilidad a los objetos nulos en el caso de los aprendientes

²⁸ Hasta aquí, hemos traducido todos los nombres de teorías e hipótesis al español y seguiremos haciéndolo con los nombres de los modelos e hipótesis del campo de la AL3. A partir de ahora, sin embargo, tras haber introducido nuestra traducción del nombre de cada modelo/hipótesis, nos referiremos a estos con sus nombres o siglas en inglés por ser estos los de uso más frecuente en el campo.

de L3 en comparación con el grupo de L2, lo cual tomaron como una muestra de la transferencia a partir de la L1. Sin embargo, las autoras son conscientes del hecho que no tenían ningún grupo del trío de lenguas L1 inglés–L2 tailandés–L3 chino para desambiguar entre los factores de la proximidad tipológica y el orden de adquisición.

Los estudios de Hermas (2010, 2014) presentan conclusiones similares a partir de datos de adquisición de inglés como L3 en los cuales se reporta una influencia exclusiva a partir de la L1 de los aprendientes bilingües L1 árabe/L2 francés. Hermas (2010) estudió el movimiento del verbo en inglés, mientras que Hermas (2014) se enfocó en dos propiedades del parámetro del sujeto nulo: los expletivos nulos y el orden de palabras en las oraciones declarativas. En ambos casos, Hermas reporta una influencia exclusiva no facilitadora a partir de la L1 árabe, y afirma que la proximidad tipológica no parece jugar un papel en el proceso de adquisición de L3, ya que el inglés se acerca más al francés que al árabe.

1.4.2 Hipótesis del factor del estatus de la L2

Más que un modelo de AL3, la hipótesis del factor del estatus de la L2 (*L2 Status Factor Hypothesis*) inicialmente formulada por Williams y Hammarberg (1998) se planteó para explicar los cambios de idioma (*language switches*) en el habla de una aprendiente de sueco como L3 (se trata de un estudio de caso). En las grabaciones recopiladas y analizadas en el marco del estudio, el idioma más utilizado para proveer palabras o estructuras no conocidas en la L3 era, de manera general, la L2 alemán, un idioma en el que la participante tenía un alto nivel de dominio. Según estos resultados, los autores afirmaron que la L2 tenía un estatus privilegiado en la adquisición de una L3 (junto con otros factores como el nivel de dominio, la tipología y el efecto de la lengua más reciente).²⁹

A partir de la hipótesis del factor del estatus de la L2 (*L2 Status Factor Hypothesis*), Bardel y Falk (2007) presentan nuevos datos de AL3 que demuestran transferencia hacia la L3 a partir de

²⁹ El ‘efecto de lo más reciente’ (*recency effect*) es un principio que procede del ámbito de la psicología. Según este principio, si una persona oye una lista de palabras, tenderá a recordar con mayor facilidad los últimos elementos de la lista, ya que son los que ha oído más recientemente (Jones & Goethals, 1972). Aplicado a la adquisición de lengua, este principio podría favorecer la influencia de la lengua más reciente, es decir, la lengua con la que el aprendiente ha tenido más contacto últimamente (Falk & Bardel, 2011, p. 63; Williams & Hammarberg, 1998, pp. 322-323).

la L2, nunca a partir de la L1. Además, los resultados muestran transferencia negativas en algunos casos, lo que va en contra del modelo de mejoramiento cumulativo (Flynn et al., 2004), que presentaremos en la sección 1.4.4. Estos resultados permiten a las autoras concluir que el estatus de la L2 es el factor de predicción de transferencia más determinante y que sobrepasa a todos los demás, incluyendo el factor tipológico.

De acuerdo con esta hipótesis, en el proceso de AL3, la L2 actúa como un filtro que bloquea el acceso a la L1. Falk y Bardel (2011) argumentan que una posible explicación para la no accesibilidad a la L1 podría proceder de las diferencias de adquisición entre la L1 y la L2 (pp. 61-62). Dado que las lenguas no-nativas se adquieren a menudo en un contexto formal (en la escuela), los aprendientes tienden a desarrollar más conocimientos metalingüísticos sobre su L2 que sobre su L1 (adquirida en un contexto natural). Entonces, la diferencia entre conocimientos explícitos e implícitos podría explicar el papel de la L2 en el proceso de AL3 (Falk & Bardel, 2010, p. 192). Trabajos más recientes sobre el *L2 status factor* (Bardel & Sánchez, 2017) ponen de relieve el papel de los conocimientos metalingüísticos como factor determinante en el proceso de adquisición de la L3.

1.4.3 Modelo de la primacía tipológica

En la misma línea que otros estudios sobre la transferencia en AL3, Rothman (2011) intenta dilucidar cuáles son los factores más determinantes a la hora de establecer la fuente de transferencia hacia la lengua meta cuando existe más de una posibilidad (como es el caso en AL3). Como respuesta a esta pregunta, el autor propone el modelo de la primacía tipológica (*Typological Primacy Model*, TPM). A pesar del uso de la palabra “tipológica” en su nombre, el modelo no considera directamente la proximidad tipológica de las lenguas implicadas sino la proximidad tipológica tal como la percibe el aprendiente, es decir, la proximidad psicotipológica (Kellerman, 1983). El TPM predice que la transferencia ocurre a partir de la lengua que el hablante percibe como más cercana, incluso en casos en los que esta opción no es necesariamente la más ventajosa.

El TPM siempre asume transferencia completa de la lengua seleccionada (L1 o L2), siguiendo la hipótesis de transferencia completa y del acceso completo de L2 (Schwartz & Sprouse, 1996), y rechaza todo tipo de transferencia parcial. En el TPM, el estado inicial de la L3 consiste en una copia de la lengua transferida (ya sea L1 o L2) y, a partir de este estado inicial, el aprendiente va ajustando y corrigiendo los errores de su interlengua. Si la lengua fuente y la lengua

meta coinciden en varios aspectos, el efecto puede ser facilitador. En cambio, cuantas más diferencias haya entre las lenguas, más ajustes tendrá que hacer el aprendiente para alcanzar la lengua meta. Rothman (2015, p. 184) argumenta que esta transferencia temprana se realiza por cuestiones de economía cognitiva. Una vez realizada la transferencia inicial, en vez de mantener tres lenguas activadas paralelamente, la mente puede inhibir de manera más eficaz la(s) lengua(s) no transferida(s) para disminuir la carga cognitiva y concentrarse en la L3.

Las evidencias empíricas a favor del modelo proceden de estudios que demuestran transferencia a partir de la lengua psicotipológicamente más cercana (ya sea la L1 o la L2) en tríos de lenguas como L1 inglés, L2 español y L3 francés/italiano (Rothman & Cabrelli Amaro, 2010) y L1 español/inglés, L2 español/inglés y L3 portugués brasileño (Cabrelli Amaro et al., 2015; Rothman, 2010); entre otros.

En principio, el TPM es un modelo que se aplica solo a la transferencia en el estado inicial de la L3, pero, de manera indirecta, formula predicciones para el desarrollo de la lengua en estados intermedios y avanzados. En efecto, la transferencia inicial debería tener un impacto a lo largo del desarrollo de la lengua, pero resulta complejo identificar las fuentes de transferencia inicial en el caso de hablantes de nivel más avanzado. Por ejemplo, si un hablante de L3 español muestra un dominio nativo de los sujetos nulos, no se puede saber si esta propiedad procede de una transferencia inicial positiva o si, por lo contrario, el estado inicial de la lengua contaba con la propiedad incorrectamente parametrizada (sujetos explícitos) pero, más adelante, se ha ido corrigiendo a medida que avanzaba el proceso de adquisición. Por último, como sugieren Alonso y Rothman (2016, p. 7), la evidencia a favor o en contra del modelo debería proceder de casos de adquisición en los estados iniciales de la L3.

1.4.4 Modelo de mejoramiento cumulativo

El modelo de mejoramiento cumulativo (*Cumulative Enhancement Model*, CEM) propuesto por Flynn, Foley y Vinnitskaya (2004). Según el CEM, todas las lenguas previamente adquiridas pueden influir en la adquisición de L3, y siempre de manera positiva. El modelo sugiere que, en casos de efectos negativos potenciales, la transferencia no se realiza.

La evidencia empírica para este modelo procede, por una parte, de estudios anteriores sobre la direccionalidad de las oraciones de relativo (núcleo inicial vs. núcleo final) en L2, que muestran

un efecto positivo de la L1 cuando esta comparte la misma direccionalidad que la L2 (Flynn, 1983, 1987; Flynn & Lust, 1980), y, por otra, de resultados más recientes con un grupo de L3 (Flynn et al., 2004). Los participantes de este grupo presentan una combinación de lenguas que permite comprobar la influencia de la L1 o de la L2 en las oraciones de relativo en la L3: kazajo L1 (núcleo final), ruso L2 (núcleo inicial) e inglés L3 (núcleo inicial). Los resultados muestran que los participantes se benefician de un efecto facilitador a la hora de adquirir la estructura de las oraciones de relativo en su L3 inglés, un efecto similar observado previamente en los hablantes de L1 español (núcleo inicial también en las oraciones de relativo) que aprenden inglés como L2. El estudio más reciente de Berkes & Flynn (2012) presenta conclusiones similares con otras combinaciones de lenguas (húngaro L1/L2, alemán L1/L2 e inglés L2/L3).

El CEM asume que la adquisición de una lengua no-nativa es un proceso no redundante en el que se transfiere todo lo que pueda resultar positivo para la L3 a partir de las lenguas previamente adquiridas. En este caso, se transfieren las estructuras de las relativas a partir de la L2, ya que estas estructuras corresponden a lo que se espera en la L3. Si bien los resultados presentados por las autoras son consistentes con el CEM, es necesario ampliar los estudios y analizar una mayor gama de combinaciones de lenguas y estructuras para determinar si, efectivamente, la transferencia de las lenguas previamente adquiridas es siempre positiva o si, por el contrario, los efectos identificados se pueden explicar mediante otros factores. Los argumentos para refutar las predicciones del CEM podrían proceder de estudios que demuestren una influencia negativa a partir de la L1 o la L2 en el proceso de adquisición de L3.

1.4.5 Modelo de proximidad lingüística

El modelo de proximidad lingüística (*Linguistic Proximity Model*, LPM; Westergaard et al., 2017) se construye sobre las bases del TPM y del CEM, pero se distingue del TPM principalmente por el tipo de influencia que predice y del CEM porque predice tanto efectos facilitadores como no facilitadores. El LPM propone que la adquisición de una L3 (o L_n) implica un aprendizaje incremental (propiedad por propiedad) a partir de una o de ambas lenguas previamente adquiridas. En un trabajo más reciente, Westergaard (2019) insiste en lo que llama potencial de transferencia completa (*Full transfer potential*) y argumenta en contra de la hipótesis de transferencia completa y del acceso completo de L2 (Schwartz & Sprouse, 1996) privilegiada

por el TPM.³⁰ Según el LPM, la influencia interlingüística ocurre de manera cumulativa solo cuando una propiedad presente en el *input* de la *Ln* muestra similitudes estructurales abstractas con las propiedades lingüísticas de una lengua previamente aprendida (Westergaard et al., 2017, p. 5). Para detectar estas similitudes, el aprendiente analiza el *input* a partir de los micro-indicios (*micro-cues*) disponibles. El concepto de micro-indicio se basa en estudios anteriores sobre la adquisición de L1 (Westergaard, 2009a, 2014). Según Westergaard (2019), el modelo de los micro-indicios (*Micro-cue Model*) representa una alternativa a los parámetros de la GU. Dicho modelo considera que la GU puede contener categorías, principios y restricciones, pero que no incluye parámetros sino micro-indicios, los cuales se desarrollan a partir de la interacción entre la GU y el *input* (p. 6). La gramática interna (lengua-I) de la L1 consiste, por lo tanto, en la acumulación de diferentes micro-indicios que permiten analizar de manera adecuada el *input* de la L1. De manera comparable, los aprendientes de L3 tienen acceso a un *input* dado que analizan a partir de los micro-indicios disponibles (de la L1 o la L2) y pueden configurar su L3 de la misma manera que su L1, su L2 o de una manera distinta si no hay similitudes entre las lenguas. En este último caso, el LPM predice que el aprendiente recurre a la GU para analizar el *input* y desarrollar nuevos micro-indicios propios del sistema de la L3. Las influencias facilitadoras proceden de un análisis correcto de las semejanzas entre las lenguas y las no facilitadoras son el resultado de un análisis equivocado (por anticipación y falta de *input* suficiente en la L3). Como en el caso del TPM, las influencias no facilitadoras no son irreversibles y el desarrollo de la L3 puede corregir las transferencias equivocadas.

1.4.6 Modelo de competición

Como mencionamos anteriormente, la adquisición de L3 ha sido estudiada sobre todo desde la perspectiva generativa y, por lo tanto, los modelos que hemos presentado hasta aquí se adscriben a esta perspectiva. Por su parte, el estudio de Sanz, Park y Lado (2015) es uno de los pocos ejemplos basados en la perspectiva cognitivo-funcionalista que miran a la adquisición de una L3 a través de la lente del modelo de competición (*Competition Model*, CM; E. Bates & MacWhinney, 1987; MacWhinney, 1987, 1997, 2008, 2012; entre otros), un modelo hasta

³⁰ El Potencial de transferencia completa se refiere a todo lo que “puede” transferirse, aclarando que no todo “debe” transferirse.

entonces solo aplicado a la adquisición de L1 y de L2. Este modelo se diferencia de los demás, entre otras cosas, porque se interesa por la adquisición de L3 desde una perspectiva cognitivo-funcionalista, pero también porque sus predicciones no se limitan a la morfosintaxis, sino que conciernen a todas las áreas de la lengua (MacWhinney, 2008).

El CM se basa en la percepción de indicios (*cues*) en la lengua para establecer conexiones entre la forma y la función. Por ejemplo, en la oración “los niños persiguen la pelota”, los dos sustantivos (niños y pelota) son dos posibles candidatos para el papel temático de sujeto. Sin embargo, en la jerarquía de los indicios del español (Kail & Charvillat, 1988), la concordancia sujeto-verbo, el orden de palabra SVO (no marcado) y el rasgo de animacidad son tres indicios que apuntan en la misma dirección, es decir, el sujeto debería ser “los niños”. La fuerza relativa de esos indicios en una lengua dada se establece con el uso y el procesamiento del *input*. Por lo tanto, un niño expuesto a otra lengua en la que la concordancia sujeto-verbo no es tan evidente (el inglés, por ejemplo) no tendrá la misma jerarquía de indicios para identificar un sujeto.

A la hora de adquirir una L2, el CM predice que los indicios de la L1 son los que determinan el procesamiento de la L2, al menos en las etapas iniciales. Luego, con el uso, se desarrolla una jerarquía propia para procesar la L2. En este sentido, el modelo predice una transferencia importante hacia la L2 en las etapas iniciales. Precisamente, Sanz, Park y Lado (2015) resumen los cuatro procesos que pueden operar a lo largo del desarrollo de la L2. La transferencia anticipatoria (*forward transfer*, de la L1 a la L2) ocurre a menudo en las etapas iniciales porque se trata de la única opción posible. Al contrario, la transferencia regresiva (*backward transfer*, de la L2 a la L1) puede ocurrir cuando los indicios de procesamiento de la L2 son utilizados para procesar la L1, un fenómeno que suele ocurrir con los hablantes cuya L2 llega a desarrollarse más que la L1, como en el caso de ciertos hablantes de herencia (Schmid, 2013) o hablantes de L2 de nivel avanzado y bilingües equilibrados (Hernandez et al., 1994; Liu et al., 1992). El proceso de diferenciación (la L1 considerada como diferente de la L2) representa otra posibilidad de influencia, pero esta vez consiste en la no transferencia. Este proceso suele ocurrir en los hablantes de L2 de nivel más avanzado (MacWhinney, 1987) y permite corregir una transferencia de la L1 a la L2 que resulta incorrecta. En este caso, con más *input*, el aprendiente reconoce nuevos indicios propios de la L2 que le permiten desarrollar una representación independiente de una estructura dada en su L2. El último proceso consiste en la amalgamación (Hernandez et al., 1994) y corresponde a la combinación de los indicios de la L1 y de la L2 para procesar las dos lenguas de

manera a la vez diferente de la L1 y de la L2. Si bien estos cuatro procesos han sido considerados principalmente en el contexto de la adquisición de L2, MacWhinney (2012) considera que se pueden aplicar de manera más amplia a la adquisición de lenguas adicionales. En principio, para establecer la transferencia hacia la L3, deberían predominar los indicios más frecuentemente activados (los de la L1).

Con esto en mente, Sanz, Park y Lado (2015) diseñaron un estudio para comprobar el modelo con hablantes de inglés L1 aprendientes de español o japonés como L2 y latín L3. Con esta triada de lenguas, los autores intentaron averiguar si la interpretación de los papeles temáticos en la L3 difería en función de la L2 de los participantes. Para responder a esa pregunta, hicieron la prueba *ab initio* y dos veces después, en un período de cuatro semanas. En la primera prueba, todos los participantes se comportaron de manera similar, mostrando una influencia de la L1 inglés (el indicio de orden SVO). En las segunda y tercera recogida de datos, si bien la diferencia entre los grupos no era significativa, se observó una ligera diferencia en función de la L2. El grupo de L2 español parecía prestar más atención a la concordancia (un indicio importante en la jerarquía de indicios del español), mientras que el grupo L2 japonés prestaba más atención al caso (más alto en la jerarquía de indicios del japonés). Las autoras argumentan que el CM predice una influencia inicial de la L1 o, por lo menos, de la lengua dominante del hablante y una posterior influencia de la L2 a medida que van recibiendo más input en la L3.

En resumen, este modelo, como otros antes mencionados, plantea la posibilidad de una influencia cumulativa y dinámica a lo largo del proceso de la L3. Si bien parece dominar la influencia de la L1 en las etapas iniciales, no excluye una posterior influencia de la L2.

1.4.7 Un modelo basado en el uso

Si bien el modelo de Competición de MacWhinney (2012) se acerca a la perspectiva basada en el uso, no toma en cuenta factores propios de esa perspectiva, como la intención comunicativa, que juega un papel importante en la adquisición del lenguaje según Tomasello (2003, 2009). Que sepamos, no existe ningún modelo basado en el uso propio de la adquisición de L3. Sin embargo, podemos extrapolar y considerar que los mismos principios que se aplican a la adquisición de L1 y L2 se aplican también a la L3, con la diferencia de que las fuentes para establecer las similitudes entre las construcciones son más diversas y que la experiencia del aprendiente está más diversificada. Un modelo basado en esta perspectiva no podría proporcionar hipótesis tan directas

como hacen los modelos presentados antes, pero las ideas que subyacen a las perspectivas basadas en el uso permiten insistir en factores menos considerados en los modelos anteriores, como la frecuencia de las construcciones presentes en el *input* y la experiencia del aprendiente.

A modo recapitulativo, resumimos en la Tabla 1.2 las principales características que definen los modelos de L3 presentados hasta aquí.

Tabla 1.2. Resumen de las principales características de los modelos de AL3

Factores	<i>L1 factor</i>	<i>L2 status factor</i>	TPM	CEM	LPM	CM
Orden de adquisición	Sí, L1 más importante	Sí, L2 más importante	No determinante	No determinante	No determinante	No determinante
Proximidad	Tipológica (no determinante)	Tipológica (no determinante)	Tipológica (más importante)	Estructural	Estructural	Estructural
Uso reciente	-	Cierta importancia	No determinante	-	No determinante	Cierta importancia
Frecuencia de uso	-	Cierta importancia	No determinante	-	No determinante	Importante
<i>Input</i>	-	Menor importancia	Cierta importancia	-	Importante	Importante
Nivel de dominio	-	Cierta importancia	Cierta importancia	-	Cierta importancia	Importante
Tipo de influencia	Solo a partir de la L1	Solo a partir de las L2 (u otras LNN)	Transferencia total (L1 o L2), en las etapas iniciales	Cumulativa, transferencia propiedad por propiedad (L1 y L2)	Cumulativa, transferencia propiedad por propiedad (L1 y L2)	Cumulativa, propiedad por propiedad a partir de la lengua dominante
Tipo de efecto	Facilitador o no facilitador	Facilitador o no facilitador	Facilitador o no facilitador	Facilitador o neutro	Facilitador o no facilitador	Facilitador o no facilitador

1.4.8 Hipótesis generales con respecto a los modelos de AL3

A lo largo de la sección 1.4, hemos presentado algunos de los modelos de AL3 más relevantes, que resumimos en la Tabla 1.2. Como se puede observar, en cuanto a los factores que determinan la(s) fuente(s) de influencia, los modelos se diferencian sobre todo en la importancia que otorgan al orden de adquisición y al tipo de proximidad (ya sea tipológica o estructural). En lo que respecta a la proximidad, los modelos que consideran la tipología como un factor principal sitúan la influencia a un nivel más global (transferencia total, a partir de una sola fuente), mientras que los que consideran las estructuras la sitúan la influencia a nivel de las propiedades lingüísticas. Otra diferencia importante entre los modelos tiene que ver con el efecto de la influencia, es decir, facilitador, neutro o no facilitador.

De cada modelo emanan ciertas hipótesis generales para nuestro estudio. A continuación, tomando en cuenta las particularidades de cada modelo, las lenguas implicadas (francés, euskera y español) y los fenómenos de la lengua (el género gramatical y los compuestos) que evaluamos en el marco de esta tesis, exponemos dichas hipótesis. Posteriormente, en los capítulos 2 y 3, tras la presentación de los fenómenos estudiados de manera más detallada, volveremos sobre estas hipótesis para precisarlas y matizarlas cuando sea necesario. De manera preliminar, considerando las características gramaticales de las lenguas de nuestro estudio, podemos considerar que una influencia del francés sería facilitadora en cuanto a los dos fenómenos, y una influencia del euskera sería neutra en cuanto al género gramatical y no facilitadora en cuanto a los compuestos.

El modelo de transferencia absoluta a partir de la L1 (*L1 factor*) predice, por un lado, un efecto facilitador de la L1 para los hablantes de francés como L1 en el caso del género gramatical y los compuestos y, por el otro, un efecto inhibitorio para los hablantes de euskera L1. En cuanto al *L2 status factor*, las predicciones serían una influencia facilitadora a partir del francés para los hablantes de francés L1 y el efecto contrario para los hablantes de euskera L1. A pesar de que todos nuestros participantes tienen las dos mismas primeras lenguas, estos dos modelos predicen resultados distintos para los dos grupos en función del orden de adquisición de cada lengua.

En cuanto al TPM, se debería observar una misma influencia facilitadora con todos los participantes, ya que todos tienen el francés (la lengua más cercana) como fuente posible de influencia para el género gramatical y los compuestos. En cuanto al LPM, se espera el mismo

resultado, ya que los micro indicios disponibles en cuanto al género gramatical y a los compuestos deberían llevar a los participantes a concluir que el francés representa la mejor opción de transferencia. El CEM también predice un efecto similar, ya que la influencia, procedente del francés, debería resultar facilitadora.

Cabe mencionar que los fenómenos estudiados en nuestra tesis se encuentran en la encrucijada entre la morfofonología, la morfosintaxis y el léxico y que los modelos considerados hasta aquí en nuestras hipótesis se han utilizado sobre todo para explicar la adquisición de la morfosintaxis en la L3. Por lo tanto, tendremos que matizar estas predicciones a la hora de hablar de cada fenómeno. Sin embargo, autores como Cabrelli Amaro (2012) han aplicado esos modelos al estudio de la fonología en L3 y abogan por la utilización de esos modelos de manera más amplia para comprobar su validez en otras áreas de la lengua.

Por su parte, el modelo de Competición tiene un alcance más amplio y puede aplicarse a todas las áreas de la lengua (MacWhinney, 2005). Este modelo predice una transferencia inicial a partir de la lengua dominante, estructura por estructura. El resultado de las predicciones es, por lo tanto, equivalente al modelo de transferencia a partir de la L1 (*L1 factor*), pero es posible que ocurra una posterior influencia de la L2. Además, como el modelo de Competición no solo se ha utilizado para predecir la influencia morfosintáctica sino también a nivel de léxico, la influencia inicial que predice debería ocurrir tanto a nivel de la morfosintaxis como a nivel del léxico.

1.5 Conclusiones

A lo largo del presente capítulo, presentamos las bases teóricas de nuestra tesis, en primer lugar, introduciendo conceptos claves, como el momento y el contexto de adquisición, el tipo de conocimiento (implícito/explicito), la dominancia, el nivel de dominio, el uso, la cronología, la multicompetencia y la influencia interlingüística y, en segundo lugar, insistiendo en las particularidades de la adquisición de L3 frente a la adquisición de L2 y de L1. Por una parte, la adquisición de una LNN presenta ciertas características que la diferencian de la adquisición de una LN; por otra parte, la adquisición de L2 y de L3 conllevan suficientes diferencias entre sí como para constituir objetos de estudio distintos. En la última parte del capítulo, presentamos las características propias de este reciente campo de estudio que es la adquisición de L3, insistiendo en conceptos claves que ya habíamos presentado antes, como la edad, el contexto y el orden de

adquisición, pero también presentando nuevas nociones como el conocimiento metalingüístico, la distancia (pisco)tipológica y estructural y el estatus particular de la L2 y la percepción de extranjería. Con base en esas nociones, presentamos los modelos más utilizados en el campo y propusimos una serie de predicciones que ofrece cada uno de esos modelos en cuanto a los fenómenos y participantes que constituyen el objeto de estudio de la presente tesis.

En los próximos capítulos, veremos cómo lo que hemos presentado en este capítulo se aplica a dos fenómenos gramaticales concretos: el género gramatical y la formación de compuestos.

*“[...] the mastery of gender is perhaps the most difficult and frustrating feature of the study of [...] a second language, especially for those whose native language lacks gender distinctions.”
(Tucker et al. 1968, p. 312)*

2 El género gramatical

La materialización del género gramatical difiere de una lengua a otra. Mientras que algunos idiomas como el inglés y el euskera carecen de un sistema completo de género gramatical,³¹ idiomas como el español y el francés tienen dos categorías de género (masculino/femenino) y otros idiomas cuentan con tres (masculino, femenino y neutro en alemán y ruso; entre otros) o más (Corbett, 1991).³² Para los hablantes nativos de una lengua con género, se trata de una categoría que no suele presentar problemas. Sin embargo, como ilustra la cita de Tucker et al. que sirve de introducción al capítulo, en adquisición de L2 el género representa un verdadero desafío. Con esta cita incompleta de la cual hemos quitado el nombre de la L2 (el francés), queremos subrayar el lugar común que representa la dificultad del género gramatical en L2. De hecho, hubiéramos podido reemplazarla por otra cita casi idéntica de Alarcón sobre el español: “the mastery of [...] gender is one of the most frustrating and difficult aspects of acquiring Spanish as a second language” (2011, p. 332; 2005, p. 2). Como se puede notar, no seremos los primeros en abordar el tema, pero, en nuestro caso, lo haremos desde el ángulo de la adquisición de L3, apoyándonos en la abundante literatura sobre el tema, tanto desde un punto de vista descriptivo como desde la adquisición y el procesamiento en lenguas nativas y no nativas.

En el marco de nuestro estudio, el género gramatical es de particular interés, ya que una de las lenguas de nuestros participantes (el francés) materializa el género gramatical del mismo modo que lo hace su L3 (el español), mientras que la otra lengua del trío (el euskera) no tiene género gramatical.

Hasta aquí, a modo de breve introducción, hemos proporcionado comparaciones entre lenguas que materializan el género gramatical con una definición intuitiva que se basa en categorías y en los géneros más habituales, sin definir lo que es el género gramatical de manera más amplia.

³¹ El inglés conserva el género solo en los pronombres personales de tercera persona del singular *he* (masculino) y *she* (femenino), mientras que el euskera lo conserva en la flexión verbal de la segunda persona del singular (tratamiento informal) con los morfemas *-k* (masculino), *-n* (femenino). Veremos más detalles sobre el euskera en la sección 2.2.3.

³² “The number of genders is not limited to three: four is common and twenty is possible.” (Corbett, 1991, p. 5).

Para adentrarnos en el tema de manera más precisa, proponemos una primera sección de nociones preliminares en la que repasaremos varias definiciones del género, contrastaremos las nociones de asignación y concordancia e introduciremos una serie de términos clave que nos permitirán sentar las bases teóricas de este capítulo. En la segunda parte, describiremos cómo el género gramatical se materializa o no en las tres lenguas de nuestro estudio. En la tercera parte, presentaremos diferentes propuestas de análisis morfosintácticos. En la cuarta parte abordaremos los estudios sobre la adquisición y el procesamiento del género gramatical en lengua nativa y no nativa. A continuación, resumiremos la información de estas cuatro primeras partes y destacaremos los puntos más importantes que guiaron el diseño del presente estudio. El diseño del estudio aparecerá en la sección 2.6, seguida por los resultados en la sección 2.7, y una discusión de los resultados en la sección 2.8 para concluir el capítulo.

2.1 Nociones preliminares

2.1.1 Hacia una definición del género

En esta sección, repasamos algunas definiciones de género, que cambian ligeramente según la perspectiva que se adopte. Desde un punto de vista muy general, Hockett (1958) afirma que “los géneros son clases de sustantivos que se reflejan en el comportamiento de las palabras que acompañan al sustantivo” (p. 231; traducción nuestra). De esta definición sobresalen tres elementos: el género está intrínsecamente relacionado a los sustantivos; sirve para clasificarlos, y tiene impacto sobre las palabras asociadas a ellos. A pesar de ser muy general, esta definición corresponde a grandes rasgos a lo que entendemos aquí por género, pero le falta precisión sobre la manera de materializar el impacto que tiene el género sobre los elementos relacionados con los sustantivos y sobre cómo se definen las clases de género. Para profundizar en el impacto del género sobre los demás elementos de una oración, es preciso hablar de concordancia. Como punto de partida, adoptamos la definición general en (2.1) citada en Corbett (1991):

- (2.1) El término *concordancia* hace comúnmente referencia a cierta covariación sistemática entre las propiedades semánticas o formales de un elemento y las propiedades formales de otro. Por ejemplo, los adjetivos pueden tomar ciertas indicaciones formales de número y género del nombre que modifican (Steele, 1978, p. 610; traducción nuestra).

Si bien esta definición resulta satisfactoria de manera general, se queda en la superficie en cuanto a cómo ocurre esa “covariación sistemática” y necesitaremos más detalles concretos sobre el proceso de concordancia, el cual exploraremos más adelante en los análisis de la sección 2.3.

De momento, continuamos con el otro elemento que faltaba en nuestra definición inicial: los detalles sobre la división en clases de género. A ese efecto, acudimos a la definición de (2.2), que se interesa, entre otras cosas, por el papel sociológico del género:

- (2.2) La palabra género [...] hace referencia al complejo fenómeno social, cultural y psicológico vinculado al sexo, un uso común de ese término en la ciencias sociales y comportamentales. La palabra género, sin embargo, también tiene un significado técnico bien establecido en el ámbito de la lingüística. El género, en este sentido más técnico consiste en una clasificación significativa de los sustantivos que tiene implicación con los fenómenos de concordancia. (McConnel-Ginet, 1988, p. 76; traducción nuestra)

Esta definición se distingue de nuestra definición inicial por el vínculo que puede tener el género con el sexo del referente, pero, al igual que en la definición de Hockett (1958), aparecen las tres nociones antes destacadas. Si trasladamos el nuevo elemento que relaciona el género con el sexo hacia un punto de vista lingüístico, podemos aludir al vínculo entre el género y el componente semántico de los sustantivos. Jarvis y Pavlenko (2008, pp. 133-135) afirman, apoyándose en el estudio de Boroditsky, Schmidt, y Phillips (2003), que existe una relación intrínseca entre género y sexo biológico (macho=masculino, hembra=femenino), de tal modo que los hablantes de una lengua con género atribuyen un sexo biológico a todos los sustantivos, aunque sea metafóricamente en el caso de los sustantivos inanimados. Esta afirmación es problemática, ya que algunas lenguas no usan ese criterio para establecer las clases de género sino otros como la animacidad (p. ej., las lenguas algonquinas; Corbett, 2006b), la humanidad o la combinación de varios criterios. Si bien parece relevante incluir la noción de semántica en nuestra definición, no podemos establecer un enlace directo y exclusivo entre sexo biológico y género. En este sentido, la definición en (2.3) es más precisa:

- (2.3) El género es:
- (i) el ordenamiento de los sustantivos en dos o más clases
 - (ii) asignado según el sexo biológico, la animacidad, y/o la humanidad, al menos *para ciertos sustantivos*
 - (iii) reflejado en los patrones de concordancia con otros elementos (p. ej., adjetivos, determinantes, verbos, auxiliares) (Kramer, 2015, p. 70; traducción y énfasis nuestros)

Esta última definición conserva los aspectos antes mencionados, pero aporta detalles sobre el papel de la semántica y va en la misma dirección que Kibort (2010), que sostiene que siempre hay una base semántica a la clasificación de género, pero que los valores de género pueden ser transparentes en mayor o menor medida (p. 77). Si bien esta definición informa sobre la asignación del género “para ciertos sustantivos” (2.3ii), no nos precisa qué ocurre en los demás casos, por ejemplo, con los asexuados, inanimados o no humanos (según la propiedad pertinente para la lengua en cuestión).

En cuanto a la asignación de género meramente gramatical para los sustantivos que no tienen un sexo biológico, investigadores como Bloomfield (1933/1973) han emitido la hipótesis de que el género se asigna de forma arbitraria, sin reglas de ningún tipo. Por su parte, Hjelmslev (1959) propuso que la clasificación por género se hacía según la “forma” y la “sustancia” de los sustantivos. Entonces, al contrario de lo que plantea Bloomfield, según Hjelmslev, no hay arbitrariedad en la asignación del género. Corbett (2006b) afirma que, en las diversas lenguas del mundo, ciertas lenguas se basan en criterios semánticos para asignar el género, mientras que otras se basan a la vez en criterios semánticos y formales, pero no hay ninguna lengua que se base exclusivamente en criterios formales (pp. 751-752). Más precisamente, en lo relativo a la asignación de género en las lenguas de nuestro estudio, existe una variedad de posturas que exploraremos en la sección 2.2 del presente capítulo para cada una de ellas (en sus subsecciones respectivas: 2.2.1 para el español, 2.2.2 para el francés y 2.2.3 para el euskera). En estas secciones describiremos también el impacto que tiene el género gramatical en los mecanismos de concordancia de cada lengua.

2.1.2 Asignación y concordancia: dos caras de la misma moneda

Las definiciones de la subsección anterior insisten en dos fenómenos relacionados con el género: la asignación y la concordancia. La asignación concierne al sustantivo y consiste en su clasificación en una u otra categoría de género gramatical. Se considera generalmente que el género gramatical es una propiedad inherente del sustantivo (S. Carroll, 1989; Carstens, 2010; Corbett, 1991). Sin embargo, esta propiedad no siempre es observable en el sustantivo mismo, sino en la concordancia con los elementos que lo modifican. Como desarrollaremos más adelante en la sección 2.2.1, el español tiene los llamados marcadores de palabras (*Word Marker*; J. W. Harris, 1991), que se asocian a lo que podría ser una marca de género (-o, -a) en el propio sustantivo, pero

la correspondencia entre marcador y género no es completamente transparente (p. ej., “una mano”, “un coche”, “una llave”). El caso del francés es aún más problemático, ya que sus terminaciones son menos transparentes. Entonces, recurrir a la concordancia para determinar el género de un sustantivo representa la única opción fiable. Por ejemplo, la única manera de confirmar que el sustantivo “nariz” es femenino en español consiste en observar las marcas de concordancia cuando aparece en combinación con un determinante (“**la** nariz”), un adjetivo (“nariz **roja**”), un participio (“su nariz está **rota**”) o un clítico (“su nariz, no **la** puedo ver más”). No obstante, para realizar la concordancia, es imprescindible conocer de antemano el género del sustantivo. Como veremos en la sección 2.4.1, en muchos estudios de adquisición de L1, se considera que la marca de concordancia de género en el determinante ayuda a los niños a asignar el género adecuado a cada sustantivo (S. Carroll, 1989; Lew-Williams & Fernald, 2007). Así, la asignación depende de la concordancia, pero no puede haber concordancia sin asignación previa de un género al sustantivo, por lo que los dos fenómenos son indisolubles.

2.2 El género gramatical en español, francés y euskera

En la presente sección, describiremos cómo se materializa el género gramatical en cada lengua del estudio, deteniéndonos en los factores que pueden contribuir a la asignación del género de un sustantivo y en el proceso de concordancia de género dentro del SD. En este sentido, exploraremos tanto la forma (morfología y fonología) como la semántica, ya que la importancia relativa de esos factores puede variar de una lengua a otra (véase p. ej., Corbett, 2006a de manera general; Poplack et al., 1982, p. 25 sobre el español)

Sea cual sea el origen del proceso de asignación de género en una lengua, de manera general, una vez se asigna el género de un sustantivo (como todavía puede ocurrir con los neologismos o los préstamos), se establece un consenso en las comunidades de hablantes y el género asignado no suele variar (Poplack et al., 1982, p. 25). Eso no impide que, para una misma lengua, el género de un sustantivo pueda variar de una comunidad de hablantes a otra. Por ejemplo, si bien el sustantivo ‘sartén’ es femenino en España y masculino en países como México, la asignación del género de este sustantivo no suele vacilar dentro de una misma zona. Se observa lo mismo en francés, por ejemplo, con el préstamo *job* ‘trabajo’, que toma el género femenino en Quebec y masculino en Francia. A pesar de las variaciones de una comunidad de hablantes a otra,

suele haber consenso sobre el uso exclusivo de un género dentro de una misma comunidad o, al menos, dentro de un mismo registro.

2.2.1 Español

El español cuenta con un 52 % de sustantivos masculinos y un 45 % de sustantivos femeninos, el resto son ambiguos en cuanto al género (Bull, 1965).³³ No existen sustantivos de género neutro en español, el cual solo aparece en algunos demostrativos (“esto”, “eso” y “aquello”), cuantificadores (“tanto”, “cuanto”, “mucho”, “poco”), pronombres personales (“ello”, “lo”) y el artículo “lo”. Por tener un alcance muy limitado, el neutro no suele ser considerado como un tercer género en español sino “el exponente de una clase gramatical de palabras que designan ciertas nociones abstractas” (Real Academia Española [RAE] & Asociación de Academias de la Lengua Española [ASALE], 2010b, p. 24).

Desde un punto de vista teórico, se considera que el género no marcado o por defecto del español es el masculino (Harris, 1991). Esta postura teórica, basada sobre todo en la mayor presencia de sustantivos masculinos en la lengua y el hecho de que la concordancia se hace con el género masculino en casos ambiguos, se ve confirmada desde el punto de vista de la adquisición (L1/L2/L3) por varios estudios que muestran una tendencia a la sobreutilización del masculino, sobre todo al inicio del proceso de adquisición (Foote, 2014; McCarthy, 2008; Montrul et al., 2008; Pérez-Pereira, 1991; Socarrás, 2012; White et al., 2004; entre otros).

2.2.1.1 Asignación

En cuanto a la asignación del género, la terminación de los sustantivos en español es, a menudo, considerada como transparente. Según el cálculo de Teschner y Russel (1984) a partir de un corpus que cuenta con casi 42 000 sustantivos con un género claramente marcado, las dos terminaciones más frecuentes son la -a (ca. 16 000 palabras) y la -o (12 552 palabras). De todas las palabras terminadas en -o, el 99,87 % son masculinas; de todas las acabadas en -a, el 96,30 % son

³³ Las cifras de Teschner y Russell (1984) apuntan hacia la misma dirección con un 52,6 % de sustantivos masculinos y 47,4 % de femeninos (quedan excluidos del cálculo los ambiguos en cuanto al género). Un cómputo que llevamos a cabo de manera semi-automatizada a partir del contenido de la última versión del diccionario de la RAE nos dio resultados similares, con un 51,68 % de sustantivos masculinos, 44,88 % femeninos y un 3,44 % que eran ambiguos, ambivalentes o epicenos.

femeninas. Visto desde otro ángulo, de todos los sustantivos masculinos, el 56,96 % termina en la vocal -o y el resto termina en -e (12,20 %), -n (9,68 %), -r (6,48 %) y tan solo un 2,69 % termina en -a.³⁴ Por el lado del femenino, la vocal -a es la terminación de un 77,50 % de los sustantivos, mientras que las terminaciones en -n y -d, representan respectivamente un 11,44 % y un 5,24 % de los sustantivos.³⁵ Apenas un 0,08 % de los sustantivos femeninos terminan en -o. Como se puede observar, en español, las terminaciones proporcionan un indicio relativamente fiable para identificar el género de un sustantivo: tan solo con sumar la proporción de sustantivos masculinos terminados en -o y los femeninos terminados en -a, llegamos a un 66,69 % de sustantivos con un género altamente previsible.

En cambio, la relación entre la semántica y el género de los sustantivos no es tan directa. Como menciona Franceschina (2005, p. 79), diferentes autores como Bergen (1978), Dinnes (1971), Da Silva (1975) y Ramsey (1956) propusieron diferentes reglas de clasificación semántica (a veces contradictorias) para los sustantivos del español. Estas son de alcance y formulación variable. Por ejemplo, Bergen propone esta simple clasificación binaria: (1) los sustantivos que se refieren a entidades animadas de sexo exclusivamente masculino son masculinos, y (2) los que se refieren a entidades animadas de sexo exclusivamente femenino son femeninos (Bergen, 1978, p. 869). Según Bergen, el género de todos los sustantivos no abarcados por esas dos reglas no se puede explicar por un criterio semántico sino fonémico. En cambio, a partir de un escrutinio de varios manuales de enseñanza del español, Dinnes (1971) enumera una lista de categorías semánticas que corresponden respectivamente al masculino y al femenino: por un lado, los machos, los números, los días de la semana, los meses del año, los ríos, los océanos, los mares, los montes, los barcos, los puntos cardinales, los países (excepto los que terminan en una -a no tónica), los infinitivos y los sustantivos de origen griego y, por el otro, las hembras y las letras del abecedario (pp. 488-489). A parte de las numerosas excepciones que implican tales reglas, la gran cantidad de reglas en sí las vuelve poco aplicables. Le parece a Franceschina (2005) poco probable que esas reglas afecten realmente la adquisición del género, ya que son establecidas, a posteriori, por lingüistas. Otro indicio sobre la menor importancia del papel de la semántica en la asignación del

³⁴ Todas juntas (y por orden de importancia), las demás terminaciones -s, -i, -z, -u, -y, -t, -x, -m, -d, -c, -j, -b, -p, -f, -k, -ch, -g y -ll representan un 6,81 % de todas las palabras masculinas.

³⁵ Todas juntas (y por orden de importancia), las demás terminaciones -s, -e, -z, -l, -i, -r, -o, -y, -x y -j representan un 5,81 % de todas las palabras femeninas.

género en español procede de datos de AL1. Según Pérez-Pereira (1991), el género de los sustantivos que se basa en un criterio semántico claro, el género natural, no se adquiere antes de los demás sustantivos. Al parecer, los niños prestan más atención a los indicios morfofonológicos (como la terminación de los sustantivos) y sintácticos (como la concordancia con los determinantes) para asociar un género a los sustantivos.

2.2.1.2 Concordancia

En español, el género de los sustantivos (y de los pronombres) tiene repercusiones en otras clases de palabras de la oración, como los determinantes, los adjetivos y los participios. Concretamente, estas partes de la oración deben concordar en género (y en número) con el sustantivo o el pronombre al que se refieren. En los siguientes ejemplos sacados del *Manual de la Nueva gramática de la lengua española* (NGLE; RAE & ASALE 2010b), se puede contemplar los distintos tipos de concordancia.³⁶

- (2.4) *la* [artículo femenino] *mesa* [sustantivo femenino] *pequeña* [adjetivo femenino];
- (2.5) *Ella* [pronombre femenino] *es muy simpática* [adjetivo femenino];
- (2.6) *este* [demostrativo masculino] *cuarto* [numeral masculino] *capítulo* [sustantivo masculino];
- (2.7) *Los* [artículo masculino] *libros* [sustantivo masculino] *eran suyos* [posesivo masculino].
(RAE & ASALE 2010b, p. 23)

De los cuatro ejemplos, nos interesa particularmente el (2.4), ya que nuestro estudio se enfoca en las concordancias entre 1) determinante y sustantivo y 2) sustantivo y adjetivo, dentro del sintagma determinante (SD).³⁷ Los cuatro ejemplos nos permiten constatar que la concordancia

³⁶ En inglés, se diferencia entre *Agree(ment)* y *Concord*, los dos se traducen como “concordancia” en español. *Agree* hace referencia a la relación y al cotejamiento de los rasgos de manera general, entre cualquier elemento de cualquier sintagma (por ejemplo, la concordancia entre sujeto y verbo), mientras que *Concord* suele hacer referencia a la relación *Agree* dentro del SD. En el marco de nuestro trabajo usaremos el término “concordancia” en todos los casos.

³⁷ Según la hipótesis del sintagma determinante (*Determiner Phrase Hypothesis*; Abney, 1987; véase Eguren, 1989 para la aplicación de la hipótesis al español), el núcleo del sintagma considerado por algunos autores como sintagma nominal (SN o *NP* en inglés) es un determinante, por lo que el sintagma ha de recibir el nombre de su núcleo. Esta hipótesis está ya asentada en el ámbito de la lingüística generativa y de aquí en adelante usaremos el acrónimo SD para referirnos a este tipo de sintagma. En algunos ejemplos de derivación sintáctica, también aparecerá la sigla DP, por coherencia con otros términos que nos abstendremos de traducir por cuestiones de comprensión por parte del lector acostumbrado a las siglas en inglés.

es bastante regular y transparente. No obstante, hay algunas excepciones a estos patrones. Como ilustra el ejemplo (2.8), en la concordancia con el determinante, el artículo femenino “la” presenta la variante “el” cuando aparece justo delante un sustantivo que comienza por una /a-/ tónica.

- (2.8) a. la famosa águila, la única aula, la escasa agua
b. el águila, el aula, el agua

Esta asimilación ante la /a/ tónica inicial del sustantivo se aplica también al artículo indeterminado “un” y los cuantificadores “algún” y “ningún”. En cuanto a la concordancia con los adjetivos, algunos adjetivos presentan variación según el género (p. ej. un libro rojo/una cama roja) y otros presentan una forma invariable (p. ej. un libro verde/una cama verde), por lo que la concordancia con el adjetivo tampoco es siempre transparente.

A pesar de esas excepciones en la concordancia de género con algunos determinantes y adjetivos en ciertos contextos precisos, el español se considera generalmente como una lengua transparente en cuanto a los indicios de género.

2.2.2 Francés

El francés, como el español, cuenta con dos géneros: el masculino y el femenino. Al igual que en español, la repartición de los dos géneros presenta un ligero desequilibrio a favor del masculino, con un 50,38 % en comparación con un 46,48 % de sustantivos femeninos y el resto con un género ambiguo (según el corpus de 9961 palabras de Lyster, 2006). Un desequilibrio más marcado aparece en el estudio de Séguin (1969), con un 58,5 % de sustantivos masculinos.³⁸ Como en español, el género considerado como no marcado o por defecto es el masculino (Ayoun, 2018; Boloh & Ibernou, 2013; Hulk & Tellier, 1999; Prodeau, 2005; Sisson, 2006).

³⁸ Estas diferencias tienen que ver sobre todo con el corpus en el cual se basa el cálculo. No tenemos los datos exactos sobre el corpus usado por Séguin (1969) ya que citamos las cifras reportadas por Ayoun (2018), pero el hecho de que ambos estudios muestren un desequilibrio en la misma dirección apunta hacia una posible replicabilidad del desequilibrio con otros corpus del francés. Cabe especificar que el corpus de Lyster contiene 9991 sustantivos inanimados que aparecen en un diccionario simplificado destinado a un público joven (*Le Robert Junior Illustré*, para los niños de 8-12 años).

2.2.2.1 Asignación

Con respecto a los factores responsables de la asignación del género en los sustantivos del francés, las reglas basadas en las terminaciones son menos transparentes que en español. Según el trabajo de Lyster (2006) que se inscribe en la línea de los trabajos de Tucker et al. (1968, 1977), seis terminaciones fonológicas predicen con un acierto de más del 90 % el género del sustantivo: /ã/ (99 %), /ɛ̃/ (98 %), /o/ (93 %) y /ɛ/ (93 %) predicen el género masculino, mientras que /z/ (97 % y /ʃ/ (90 %) predicen el femenino. Sin embargo, estas terminaciones representan apenas el 18 % de todos los sustantivos del corpus de Lyster. Para cubrir una mayor parte de su corpus y alcanzar un mayor grado de precisión en la previsibilidad de las terminaciones, Lyster recurre a la ortografía de la rima y así logra predecir el género del 80 % de los sustantivos masculinos y del 81 % de los sustantivos femeninos. No obstante, la lista de las terminaciones ortográficas que predicen el género de manera consistente cuenta con decenas de entradas, tanto para el masculino (más de 80) como para el femenino (más de 70). Por su parte, más que en la terminación ortográfica o fonológica, Meisel (2018) insiste en el papel de la morfología derivativa. Por ejemplo, las palabras terminadas por los sufijos *-age* (*plumage*, “plumaje”), *-ier* (*vinaigrier*, “vinagrero”), *-ment* (*logement*, “alojamiento”), *-isme* (*bilinguisme*, “bilingüismo”) marcan sistemáticamente el masculino, mientras que los sufijos *-tion* (*participation*, “participación”), *-té* (*bonté*, “bondad”), *-tié* (*amitié*, “amistad”) marcan sistemáticamente el femenino (Meisel, 2018, p. 659). Si bien estas reglas pueden predecir de manera eficaz el género de los sustantivos morfológicamente complejos, no sirven para los sustantivos simples que comparten la misma ortografía que algunos de esos sufijos, pero que pertenecen al género opuesto, como *page* (femenino, “página”), *dent* (femenino, “diente”), *bastion* (masculino, “bastión”), *côté* (masculino, “lado”). En resumen, la forma (ya sea fonológica, ortográfica o morfológica) es útil para establecer patrones de género, pero ninguno de éstos tiene el mismo poder predictivo que los del español.

En cuanto a la relación entre semántica y género, apenas el 10,5 % de los sustantivos en francés tienen un género semántico, es decir, basado en el género natural del referente (Ayoun, 2010; H. Séguin, 1969). Además, cuando hay un conflicto entre la forma y la semántica, la forma tiene prioridad, como en *la sentinelle* (“la centinela”) que siempre es femenina, aun cuando el referente es un hombre (Meisel, 2018, p. 659). Aparte de los sustantivos animados, cuyo género gramatical suele corresponder con el género natural, ciertos autores observan otros patrones semánticos en la repartición del género gramatical en los inanimados. Por ejemplo, Surridge (1989)

nota que ciertos grupos semánticos, como los días de la semana, los meses del año, las temporadas, las lenguas, los colores, los árboles, los nombres de vino y de queso (entre otros) son principalmente masculinos, mientras que los días festivos (*la Saint-Jean Baptiste, la Toussaint, etc.*), las disciplinas académicas y los tipos de automóviles son generalmente femeninos. Por su parte, Rice (2006) observa que los nombres que se refieren a una ruta son generalmente femeninos (con las excepciones *chemin* “camino” y *sentier* “sendero”).

En resumen, la asignación del género en francés parece vinculada a una combinación de reglas basadas en la forma y en la semántica que resultan difíciles de enumerar de manera clara y concisa, dada la complejidad de sus interacciones. Varios intentos de formalización de las reglas (Lyster, 2006; H. Séguin, 1969; Tucker et al., 1968, 1977; entre otros) han contribuido de manera considerable a precisar la descripción del fenómeno, pero, como afirma Comrie (1999), deben entenderse como racionalizaciones resultantes de análisis lingüísticos y no necesariamente como un reflejo de la mente del hablante. De hecho, dada la complejidad de estas reglas, se considera generalmente que el hablante tiene que recurrir al contexto sintáctico (concordancia con el determinante, el adjetivo y los participios) de un sustantivo para asignarle el género que le corresponde (Jakubowicz & Roulet, 2007, p. 188). En adquisición de L1, el indicio más utilizado es la concordancia entre determinante y sustantivo (S. Carroll, 1989; Sourdot, 1977).

2.2.2.2 Concordancia

La concordancia dentro del SD se hace en francés de la misma manera que en español, es decir que, el determinante y el adjetivo deben concordar en género con el sustantivo que modifican. Una de las diferencias entre el español y el francés consiste en que algunos sustantivos –*amour* “amor”, *délice* “delicia” y *orgue* “órgano”– pasan de masculino a femenino cuando aparecen en plural, por lo tanto, la concordancia en los ejemplos (2.9) es perfectamente aceptada, a pesar de ser algo excepcional y marginal.^{39,40}

(2.9) a. *Le bel amour*

³⁹ Si bien tradicionalmente se solía preferir la variante en femenino plural, hoy en día, los dos géneros son aceptados para la forma plural y se suele reservar el femenino para un uso estilístico («Amour» Office Québécois de Langue Française [OQLF], 2018)

⁴⁰ Este fenómeno también se puede observar de manera muy ocasional en español con sustantivos como arte (el arte, las artes).

DET.sg.masc bello.sg.masc amor.sg.masc
b. *Les belles amours*
DET.pl bellas.pl.fem amores.pl.fem

En cuanto a la concordancia con los determinantes, el francés muestra un repertorio de determinantes menos variable que el español en cuanto al género. En ese sentido, los determinantes no son siempre un indicio fiable de género, ya que el contraste que se establece entre masculino y femenino queda neutralizado en algunos casos. La Tabla 2.1 presenta la lista de los determinantes del francés en función del género y del número del sustantivo con el que se establece la concordancia.

Tabla 2.1. Lista de determinantes en francés

	Masc. sing.	Fem. sing.	Masc plu.	Fem. plu.
Artículos indefinidos	un	une	des	des
Artículos definidos	le	la	les	les
Artículos definidos *	l' <i>l'avion</i>	l' <i>l'erreur</i>	les <i>les avions</i>	les <i>les erreurs</i>
Demostrativos	ce	cette	ces	ces
Demostrativos *	cet	cette	ces	ces
Posesivos 1ª persona sing.	mon	ma	mes	mes
Posesivos 1ª persona sing.*	mon <i>mon avion</i>	mon <i>mon erreur</i>	mes <i>mes avions</i>	mes <i>mes erreurs</i>
Posesivos 2ª persona sing.	ton	ta	tes	tes
Posesivos 2ª persona sing.*	ton <i>ton avion</i>	ton <i>ton erreur</i>	tes <i>tes avions</i>	tes <i>tes erreurs</i>
Posesivos 3ª persona sing.	son	sa	ses	ses
Posesivos 3ª persona sing.*	son <i>son avion</i>	son <i>son erreur</i>	ses <i>ses avions</i>	ses <i>ses erreurs</i>
Posesivos 1ª persona plu.	notre	notre	nos	nos
Posesivos 2ª persona plu.	votre	votre	vos	vos
Posesivos 3ª persona plu.	leur	leur	leurs	leurs

* ante un sustantivo con vocal inicial

Como podemos observar, en varios casos no hay diferencia entre la forma masculina y la femenina e, incluso cuando la hay, esta diferencia ortográfica puede ser difícilmente perceptible en la modalidad oral con los sustantivos que empiezan por vocal. Por ejemplo, con el artículo indefinido, la diferencia entre *[ynotobus] y [œnotobys] (**une autobus*, *un autobus*, “un autobús”) resulta difícil de percibir con la ligación entre el artículo y el sustantivo, ya que esta distinción

tiende a neutralizarse en la forma [œnotobys] (Barbaud et al., 1982).⁴¹ Estos sustantivos con vocal inicial presentan un desafío adicional para los hablantes de francés como lengua no nativa (L2 o L3) (Dewaele, 2015; Granfeldt, 2005). A pesar de la neutralización del género en algunos determinantes, éstos siguen siendo los indicios más fiables para la identificación del género de un sustantivo (Jakubowicz & Roulet, 2007, p. 189).

En cuanto a la concordancia con los adjetivos, según un estudio sobre 12000 adjetivos llevado a cabo por Séguin (1973), en la modalidad oral, apenas el 33 % de los determinantes tienen una forma variable en cuanto al género. En la modalidad escrita, esta tasa sube al 58 %, pero eso significa que, aun así, queda un 42 % de adjetivos cuya concordancia no es perceptible.

El estudio llevado a cabo por Ayoun (2010) resume de manera clara la presencia de las marcas de género en los determinantes y adjetivos del francés. En un corpus constituido por 5016 SD extraídos de 34 artículos de periódicos y revistas, el 41,22 % de estos presenta marcas de género en el determinante y apenas el 9,01 % presenta marcas de género en el adjetivo. Es decir que, en casi la mitad de los SD, la concordancia no ofrece indicios sobre el género del sustantivo.

En resumen, si bien el francés comparte mucho con el español en cuanto a la asignación y concordancia de género gramatical, hay que destacar que el francés muestra una mayor opacidad tanto en lo relativo a la asignación (ausencia de terminación tan transparentes, como -o o -a) como a la concordancia (ausencia de marca de género en una variedad de determinantes y adjetivos). Como veremos en la sección 2.4, esto no parece dificultar la adquisición del género en francés como L1, pero sí presenta un gran desafío en francés como L2/L3.

2.2.3 Euskera

Como hemos mencionado en la introducción del presente capítulo, el euskera no materializa un sistema de género gramatical de la manera que lo hacen el español o el francés.

La distinction des genres grammaticaux se fait en basque d'une façon très originale. Le basque ne concorde sur ce point ni avec les langues qui lui sont le plus directement apparentées et dont il est

⁴¹ En el habla de la provincia de Quebec, el género del sustantivo “autobus” como de otros sustantivos con vocal inicial (*affaire, heure*) varía de una región a otra y de un registro a otro, como lo subraya Remysen (2003, p. 33)

séparé depuis plusieurs millénaires, ni avec celles au contact desquelles il vit depuis de nombreux siècles. (Lafon, 1947, p. 393)⁴²

De hecho, el euskera es, a menudo, considerada como una lengua sin género gramatical (Trask, 2003). Sin embargo, podemos subrayar dos muestras de género en euskera: (1) en algunos sustantivos que cambian según su género biológico y (2) en el uso de cierta forma de tratamiento de segunda persona del singular conocida como *hika*. Trataremos el primer punto a continuación y el segundo en el apartado sobre la concordancia.

2.2.3.1 Asignación

En euskera, como ocurre con la mayoría de los sustantivos animados de las dos lenguas antes mencionadas, el género se asigna a los sustantivos según su género biológico. Del mismo modo que el castellano tiene una palabra para “hombre” y otra para “mujer”, el euskera tiene las palabras “*gizon*” y “*emakume*”. Sin embargo, al contrario del castellano, que cuenta con cientos de palabras de ese tipo, estos pares de palabras son escasos en euskera. La Tabla 2.2 muestra algunos ejemplos del euskera en comparación con sus equivalentes en español.

⁴² Las lenguas “más emparentadas” a las que se refiere Lafon son las lenguas caucásicas. Esta hipótesis de parentesco gozó de cierto reconocimiento durante décadas, pero hoy en día parece poco probable (Trask, 1995).

Tabla 2.2. Ejemplos de sustantivos en euskera que tienen una forma diferente en masculino y en femenino, junto con sus equivalentes en español.

Español			Euskera		
Masc.	Fem.	Pl.	Masc.	Fem.	Pl. ⁱ
macho	hembra		arra	eme	
hombre	mujer		gizon	emakume	gizon-emakumeak
señor	señora	señores	jaun	andre	jaun-andreak
padre	madre	padres	aita	ama	gurasoak
hijo	hija	hijos	seme	alaba	seme-alabak
hermano	hermana	hermanos	anaia ⁱⁱ	arriba ⁱⁱ	anai-arrebak
tío	tía	tíos	osaba	izeba	osaba-izebak
chico	chica	chicos	mutil	neska	mutil-neskak
toro	vaca		zezena	behi	
gato	gata	gatos	katu	kateme	
perro	perra	perros	txakur arra	txakur eme	

ⁱ Los plurales colectivos del euskera hacen referencia a un grupo mixto formado por elementos de ambos géneros biológicos. En el caso contrario, se forma el plural solo con uno de los elementos. Por ejemplo, “*jaun-andreak*” equivale a “los señores”, “*jaunak*” equivale también a “los señores”, pero en un grupo solo de hombres, y “*andreak*” equivale a “las señoras”.

ⁱⁱ Existen palabras diferentes en euskera para referirse a los hermanos entre ellos, según su sexo: “*anaia*” se refiere al hermano de un hermano, “*neba*” al hermano de una hermana, “*arriba*” a la hermana de un hermano y “*ahizpa*” a la hermana de una hermana.

2.2.3.2 Concordancia

En lo referente a la concordancia de género en euskera, solo se puede hacer referencia a las formas verbales conjugadas con el pronombre “*hi*” (segunda persona informal, equivalente de “tú” en castellano o “*tu*” en francés), que cambian según si el interlocutor es de sexo masculino o femenino. El ejemplo (2.10) muestra una conjugación con el pronombre “*zu*” (equivalente de “usted” en castellano o de “*vous*” singular en francés), mientras que (2.11) muestra la conjugación con “*hi*” cuando el interlocutor es femenino y (2.11) cuando es masculino.⁴³

(2.10) (*zu-k*) *Lan egin du-zu*
 (usted-ERG) Trabajo hacer AUX-2SG.ERG
 “(usted) ha hecho el trabajo.”

⁴³ En las glosas en euskera, utilizamos el guion para separar los distintos morfemas cuando resulta pertinente para el análisis, como aquí para destacar el sufijo ergativo (*-k* o *-n* según los ejemplos). Nuestra traducción de la glosa sigue la de Echevebarria en inglés que traduce ambos ejemplos por “*You have done work*”. Una traducción menos literal y más exacta sería “*You have worked*” (has trabajado), dado que “*lan egin*” está lexicalizado como un verbo en sí y ya no se considera que significa “hacer trabajo”, sino “trabajar”.

- (2.11) a. *(hi-k) Lan egin du-n*
 (tú-ERG) Trabajo hacer AUX-2SG.FEM.ERG
 b. *(hi-k) Lan egin du-k*
 (tú-ERG) Trabajo hacer AUX-2SG.MASC.ERG
 "(tú) Has hecho el trabajo" (Echeverría, 2011, p. 276)

La forma de tratamiento *hika* va más allá de la simple concordancia verbal con el pronombre *hi*. De hecho, todos los verbos conjugados deben llevar la marca femenina o masculina correspondiente a *hika* cuando se utiliza esta forma para dirigirse a un interlocutor, fenómeno conocido como “alocutividad” (*allocutivité*, en francés, término acuñado por Bonaparte, 1862) y ampliamente documentado en los distintos dialectos del euskera, aunque asociado al ámbito rural (Alberdi, 1995; Oyharçabal, 1993 entre otros; véase también Eguren, 2009 para un uso del término en castellano).⁴⁴ El ejemplo (2.12) muestra una conjugación en primera persona del singular que lleva también el sufijo -n correspondiente al *hika* femenino.

- (2.12) *(Ni) Hirekin etorri nau-n*
 (yo) contigo venir AUX.1SG.FEM.ALO
 “(yo) He venido contigo” (Alberdi, 1995, p. 277)

Ya que la presentación del sistema verbal del euskera se sale de los objetivos del presente trabajo, no entramos en más detalles y referimos al lector a los autores antes citados para una visión más exhaustiva del *hika* en euskera.

2.2.4 Síntesis de las comparaciones entre las lenguas del estudio

El euskera difiere de las dos lenguas romances estudiadas en cuanto a la concordancia de género. Si bien podemos hallar algunas pistas vinculadas al género natural de los sustantivos animados en esta lengua, no hay nada que se acerque al concepto de género gramatical del francés y el español. Por su parte, esas dos últimas lenguas comparten las dos mismas categorías de género (masculino y femenino) y presentan mecanismos de concordancia similares con los mismos elementos (determinantes, adjetivos y participios). Sin embargo, existen diferencias entre las dos lenguas romances, la principal de las cuales reside en la mayor transparencia del género en español

⁴⁴ Eguren (2009) afirma que “[l]a expresión morfológica de la alocutividad es, por lo que se [le] alcanza, un fenómeno poco común en las lenguas del mundo.” (p. 2)

(mediante la -a y la -o en los distintos elementos que concuerdan) comparado con el francés (Jakubowicz & Roulet, 2007, p. 189).

2.3 Análisis morfosintáctico

En francés y en español se considera generalmente que el rasgo de género es inherente al sustantivo y que la concordancia con el determinante y el sustantivo resulta de una operación sintáctica en el SD. Desde la perspectiva lexicalista, los rasgos de los sustantivos están especificados directamente en el léxico (Chomsky, 1995; Müller & Wechsler, 2014), mientras que desde la perspectiva de la Morfología distribuida (*Distributed Morphology*; Marantz, 1997), las raíces léxicas adquieren sus rasgos al fusionarse con un nodo nominalizador (*n*), que comporta un paquete (*bundle*) de rasgos abstractos, como el de género (Kihm et al., 2008; Kramer, 2015). Este nodo *n* entra en la sintaxis plenamente especificado. Por ejemplo, el nodo terminal sintáctico para generar el sustantivo “chica” correspondería (de manera simplificada, solo con el rasgo de género) a $\sqrt{\text{CHIC}}-[n_{+\text{FEM}}]$, donde $[n_{+\text{FEM}}]$ especifica a la vez la categoría léxica (*n* nominaliza la raíz) y el género (+FEM).

En ambos casos, el sustantivo entra en la sintaxis con una serie de rasgos llamados *phi-features*, normalmente representados mediante la letra griega ϕ , entre los cuales se encuentran el género y el número. En algunos elementos de los sintagmas, los rasgos- ϕ son interpretables y en otros no lo son.⁴⁵ El cotejamiento entre estos rasgos interpretables y no interpretables es responsable de las operaciones de concordancia que detallaremos a continuación.

Las implicaciones de la teoría de los rasgos- ϕ son diversas, ya que abarca tanto aspectos de la sintaxis, de la semántica como de la morfología (Harbour et al., 2008). Eso permite distintos acercamientos a fenómenos como el que nos interesa aquí, la concordancia de género. Sin embargo, en la presente sección, nos limitaremos al ámbito de la sintaxis, y más precisamente, dentro del sintagma determinante (SD). Las operaciones de concordancia de género dentro del SD están estrechamente vinculadas al rasgo de número, por lo que varios estudios se interesan por los dos

⁴⁵ Según la Teoría del cotejamiento (*Checking Theory*) del Programa minimalista, todos los rasgos no interpretables deben ser eliminados en la Forma Lógica (FL; *Logical Form*) antes de pasar a la Forma Fonológica (FF; *Phonological Form*). Para ser eliminados, tienen que ser cotejados. Para más información sobre esa teoría, referimos al lector a los trabajos de referencia de Chomsky (1995, 1999, 2000).

fenómenos a la vez (véase por ejemplo Acuña-Fariña et al., 2014; Barber & Carreiras, 2005; Dowens et al., 2011; Faussart et al., 1999; Picallo, 2008; Sagarra & Herschensohn, 2010; White et al., 2004). Si bien ambos rasgos tienen un impacto en la sintaxis del SD, difieren en su esencia: el rasgo de género es inherente al sustantivo y puede no tener ningún valor semántico, mientras que el rasgo de número no es inherente al sustantivo y tiene una interpretación semántica. Por su naturaleza y por la posible influencia que puede ocurrir a partir del francés (pero no a partir del euskera), en el marco de nuestro estudio nos interesamos únicamente por el rasgo de género y, por lo tanto, obviamos los detalles sobre el impacto del rasgo de número en las operaciones sintácticas.

Según Carstens (2000, 2001; tomado de Franceschina, 2005, pp. 85-89), la concordancia se define de la siguiente manera:

(2.13) En la relación de concordancia:⁴⁶

- a. Una sonda (*probe*) α tiene rasgos- ϕ no interpretables.
- b. Una meta (*goal*) β tiene los correspondientes rasgos- ϕ
- c. Los rasgos- ϕ no interpretables son cotejados y eliminados
- d. Si α tiene rasgos intrínsecos de Caso valorados, coteja los rasgos de Caso sin valores de β ; después, ambos rasgos de Caso son eliminados. (Carstens, 2001, p. 151; traducción nuestra)

La etapa de (2.13d) tiene implicaciones fuera del SD, por lo que nos concentraremos en las etapas (2.13a-c).⁴⁷ Concretamente, dentro de un SD, el sustantivo funciona como meta y el determinante y el adjetivo como sondas. Éstos tienen que cotejar sus rasgos no interpretables [UGEN] con los valores correspondientes del sustantivo, lo que desencadena las operaciones de concordancia y, en ciertos casos, de movimiento.⁴⁸ En el ejemplo (2.14a), el rasgo no interpretable [UGEN] del determinante “el” recibe su valor MASC del rasgo interpretable [GEN:MASC] del

⁴⁶ En inglés, se utiliza a veces el término *Agree(ment)* y otras veces *Concord*, que traducimos en ambos casos por “Concordancia” en español. *Agree* hace referencia a la relación y al cotejamiento de los rasgos de manera general entre cualquier elemento de cualquier sintagma (por ejemplo, la concordancia entre sujeto y verbo), mientras que *Concord* suele hacer referencia a la relación *Agree* dentro del SD. En el marco de nuestro trabajo usaremos el término “Concordancia” en todos los casos y especificaremos el tipo de relación de concordancia cuando sea necesario.

⁴⁷ Las reglas de concordancia se aplican también fuera del SD, por ejemplo, entre sujeto y atributo en “La-FEM mesa-FEM es redonda-FEM”. Sin embargo, dado que el presente trabajo se centra en las operaciones sintácticas dentro del SD, no contemplaremos las reglas más allá del SD al quedarse estas fuera del alcance de nuestro estudio.

⁴⁸ Prefijamos los rasgos con la letra u cuando no son interpretables (del inglés *uninterpretable*).

sustantivo “lápiz”. Lo mismo ocurre con el adjetivo y el determinante del ejemplo (2.14b), pero esta vez con el valor FEM.

- (2.14) a. El[UGEN:MASC] lápiz[GEN:MASC]
b. La[UGEN:FEM] cama[GEN:FEM] roja[UGEN:FEM]

En resumen, la operación de concordancia depende del género asignado al sustantivo, que es un rasgo inherente. La teoría expuesta hasta aquí explica la concordancia de género en los casos en los que el hablante tiene un género claramente asignado a cada sustantivo (como suele ser el caso en la L1), pero en el proceso de adquisición de una lengua no nativa, o en situaciones de alternancia de código (como veremos en la sección 2.4.2), no queda claro si cada sustantivo tiene un género asignado. Por ejemplo, un hablante no nativo del español bien podría no tener un género asignado para el sustantivo “lápiz”, como en (2.15a), o tener el género equivocado, como en (2.15b).

- (2.15) a. El/*la [] lápiz[]
b. La[uGEN:FEM] *lápiz[GEN:FEM]

Para el caso de (2.15a), la propuesta de Carstens (2000, 2001) no nos permite predecir lo que pasa cuando la sonda o la meta (o ambos) están ausentes de la computación. Para el ejemplo (2.15b), damos por hecho que el género atribuido al sustantivo lápiz es erróneo, pero también podríamos explicar el error en superficie por un problema en el proceso de concordancia. En la sección 2.4.2, cuando hablemos de alternancia de código en los bilingües, volveremos sobre los problemas de asignación y concordancia del género dentro del SD, con una propuesta de Liceras et al. (2008).

El análisis morfosintáctico propuesto hasta aquí sobre la concordancia dentro del SD se adscribe a la perspectiva generativista y ha sido aplicado de manera similar a otras lenguas romances, como el francés, que, como vimos en la sección anterior, presenta los mismos rasgos de género y establece la concordancia con los mismos elementos dentro del SD.

Desde una perspectiva constructivista, más en línea con los acercamientos basados en el uso, Mariscal (2009) considera que el sistema de género se construye de manera gradual y que no está guiado por reglas y rasgos formales de género. La adquisición del género está relacionada con la frecuencia de coocurrencia de las diferentes formas de determinantes y adjetivos con los

sustantivos. Estos patrones son a la vez los que guían la asignación del género a cada sustantivo y los que permiten el desarrollo de las estructuras abstractas del tipo artículo + sustantivo que presentan concordancia de género. A partir de esta estructura básica de concordancia, el fenómeno se generaliza a otras estructuras y la concordancia se puede realizar con los adjetivos dentro del SD, los adjetivos en el predicado, etc. Que sepamos, no existen análisis morfosintácticos detallados de la concordancia de género en los acercamientos basados en el uso, ya que no es el foco de esta perspectiva.

2.4 Adquisición y procesamiento del género gramatical

En la L1, se considera que los hablantes nativos de un idioma con género gramatical llegan a dominar este fenómeno en torno a los tres o cuatro años (Montrul, 2004) y los errores son escasos a lo largo del proceso de adquisición, como demuestran numerosos estudios sobre el francés (Clark, 1985; Grégoire, 1947; Kupisch et al., 2002; Patsy Martin Lightbown, 1977) y el español (Hernández Pina, 1984; Pérez-Pereira, 1991; Socarrás, 2012). En cambio, como menciona Foote (2014, p. 365), numerosos estudios en el campo de AL2 sugieren que el género gramatical representa un desafío para los aprendientes de L2 y que esta dificultad puede perdurar hasta los niveles avanzados de dominio de la lengua no nativa (véase Franceschina, 2001, 2005; Grüter et al., 2012; entre otros). A continuación, repasamos estudios sobre la adquisición del fenómeno en contextos de L1, 2L1, L2 y L3.

2.4.1 En L1

Si bien la adquisición del género en una L1 culmina generalmente con la ausencia casi completa de errores, los pocos errores que aparecen a lo largo del proceso y las etapas por las que pasan los niños nos informan sobre las dificultades que subyacen al fenómeno y, de este modo, ofrecen pistas para explorar su adquisición en una lengua no-nativa. La información disponible sobre el desarrollo del fenómeno en español L1 procede de datos longitudinales recopilados en contexto natural (Hernández Pina, 1984; López Ornat, 1997; Snyder, 1995), en contexto experimental (Pérez-Pereira, 1991; Socarrás, 2012) o de una mezcla de datos espontáneos y elicitados (López Ornat, 1997, 2003; López Ornat et al., 1994)

En nuestra anterior descripción del fenómeno en español y en francés, insistimos en los indicios que guían la asignación del género a los sustantivos desde un punto de vista meramente

descriptivo: el género por defecto, la morfofonología, la semántica y la concordancia. A continuación, exploramos el impacto que parecen tener estos indicios en adquisición de L1.

Hernández Pina (1984), documenta el proceso de adquisición del género en español L1 por parte de su hijo (Rafael) hasta los tres años.⁴⁹ Si bien el niño empieza a producir marcas del género en los determinantes hacia los 18-19 meses de edad, la autora documenta errores de no concordancia como “un llave”, “un leche”, “una camión” o “una pez”. Luego, Rafael pasa por una etapa en la que sobregeneraliza el indicio de género tanto en sustantivos (p. ej., “una mota rota”) como en adjetivos (p. ej., “la tierra azul”). En la producción del niño, la concordancia con los determinantes precede a la concordancia con los adjetivos. A los 2;7 años, los errores de concordancia con los determinantes son escasos.

En López Ornat (1994, 1997), los datos longitudinales de una niña (María) entre 1;7 y 4;0 años muestran distintos pasos en la evolución del sintagma nominal, empezando con una gran mayoría de sustantivos solos, sin determinantes, para luego pasar por una etapa transicional del llamado protodeterminante o relleno monosilábico (*Monosyllabic Place Holder*; Bottari et al., 1993) en la que se produce una vocal (a, e, o, u) antes del sustantivo y, posteriormente, culminar en la producción de secuencias determinante-sustantivo en las que el determinante generalmente concuerda en género (pero no siempre en número) con el sustantivo.

En Schnell de Acedo (1994), los datos espontáneos de una niña procedentes de tres grabaciones distintas (a los 16, 23, y 29 meses de edad) muestran que la producción de determinantes evoluciona de manera gradual en el tiempo. En la primera etapa, los pocos determinantes producidos son casi exclusivamente masculinos y en las dos etapas siguientes aparecen los determinantes femeninos. La autora observa una sola producción de un artículo indefinido en la primera etapa y, si bien no concuerda con el sustantivo (“una papá”, en el sentido de padre), este error muestra que la niña parece prestar atención a la terminación en -a del sustantivo. Aparte de este error, los datos de Schnell de Acedo (1994) presentan una producción casi ausente de errores de concordancia.

⁴⁹ Como menciona Socarrás (2012), desafortunadamente, los datos presentados por Hernández-Pina (1984) son solo descriptivos, por lo que resulta difícil evaluar hasta qué punto los errores de género reportados son frecuentes en la lengua del niño.

El estudio de Pérez-Pereira con 160 niños de entre 4 y 11 años revela que el género semántico (sexo) parece influir poco en la asignación del género, mientras que la información morfofonológica (la terminación) y sintáctica (la concordancia) tienen más peso. En su experimento, en el cual los niños tenían que asignar un género a unos sustantivos inventados según diferentes indicios sintácticos (p. ej, “**un** tanten”, “**una** pilin”, o “dos palires”), morfofonológicos (p. ej., “dos carepos”, “dos aniras”, “dos palires”) y semánticos (representados por un indicio de género natural en la imagen para los animados masculinos o femeninos o sin indicios para los inanimados). El autor observó una tendencia hacia la asignación más frecuente del género masculino, sobre todo cuando había conflicto entre los tres tipos de indicio, de manera incluso más pronunciada que en un estudio comparable en francés (Karmiloff-Smith, 1979). El indicio sintáctico (el género del determinante) parece ser el más fiable para los niños, ya que, en casos de conflicto, se privilegia este indicio. Este efecto se observa desde los 4 años de edad y parece afianzarse de manera más temprana que en francés (Karmiloff-Smith, 1979). A su vez, el indicio morfofonológico es el segundo más relevante para los niños y, en último lugar, aparece el indicio semántico que solo es usado cuando no hay conflicto entre este y los indicios sintáctico (determinante plural sin marca de género) y morfofonológico (terminación opaca). En cuanto al tipo de concordancia, el autor observa solo un 3,6 % de respuestas en las que los niños producen un adjetivo que no concuerda con el género del resto del SD (p. ej. “la linola blanco”), las cuales se dan sobre todo en los niños de 4 y 5 años. Esto parece indicar que, si bien los niños empiezan a hacer la concordancia con el determinante antes que con el adjetivo (Hernández Pina, 1984), el mecanismo de concordancia como tal parece estar ya estabilizado a los 4 años.

En el estudio de Mariscal (2009) con cuatro niños de entre 1;10 y 2;01 años, la autora recogió datos espontáneos y elicitados, a lo largo de dos o tres ciclos según el progreso de cada niño.⁵⁰ Una parte de los datos elicitados procede de una tarea que se diseñó específicamente para fomentar la producción de adjetivos, ya que se había observado en estudios anteriores que los adjetivos aparecían de manera escasa en los datos espontáneos de niños de estas edades. En cuanto a la producción de artículos antes del sustantivo, la autora nota que todos los niños pasan por una etapa en la cual producen el sustantivo solo, sin artículo; luego, tres de los cuatro niños producen un relleno monosilábico en lugar de un artículo completo; después todos pasan a una fase en la

⁵⁰ Cada ciclo está formado por varias sesiones de grabaciones.

cual la producción oscila entre la secuencia artículo (le, la, un, una) + sustantivo y otro determinante (este, otro, etc.) + sustantivo. En cuanto a los adjetivos, incluso con la tarea de elicitación, la producción fue escasa, pero suficiente para permitir a la autora abstraer una secuencia de adquisición: producción de una unidad no analizada > productividad mínima > generalizaciones y errores > ausencia de errores. De la etapa de generalizaciones y errores, la autora destaca el hecho de que la mayoría de los errores de concordancia implican un adjetivo acabado en -o que modifica un sustantivo femenino, algo que casi no ocurre a la inversa. La autora también observa que la desaparición de errores de concordancia en los adjetivos parece ir a la par con la estabilización en la producción de los determinantes, es decir, cuando los niños producían menos de un 10 % de omisión de determinantes. La autora, que adopta una perspectiva cognitivista, concluye que la generalización de la estructura X + N (donde X es generalmente un determinante) representa el paso previo a la adquisición de la morfología de género. El análisis de la forma del determinante, del sustantivo y del adjetivo, así como su distribución, representan la clave de la aparición gradual de la concordancia de género en el sistema lingüístico del niño.

En otro estudio experimental, Socarrás (2012) utilizó una marioneta con diferentes juguetes y objetos para elicitación de sintagmas determinantes con tres niños de menos de tres años (Elián 2;1, Alonso 2;2 y Londa 2;9) en dos o tres sesiones individuales para cada uno. La autora también recogió datos comparables con dos niños de más de tres años para contrastar el nivel de dominio de la concordancia de cada grupo. Los resultados muestran que los errores de concordancia con los determinantes son escasos y, cuando los niños producen solo una vocal en lugar de un determinante, esa vocal también concuerda generalmente con el sustantivo. El papel de la morfofonología no parece influir en la concordancia, ya que los niños hacen la concordancia tanto cuando la terminación es transparente como cuando es opaca. Además, en los pocos errores producidos, los niños muestran una preferencia por el género masculino por defecto, es decir que usan un determinante en masculino con un sustantivo femenino con más frecuencia que un determinante femenino con un sustantivo masculino. En cuanto a la concordancia con los adjetivos, los dos niños más jóvenes producen menos adjetivos, mientras que la niña produce más, pero con una tasa de error importante (40 %).

En resumen, en el proceso de adquisición de español como L1, los estudios indican que 1) la concordancia entre determinante y sustantivo precede a la concordancia entre sustantivo y adjetivo, 2) el masculino es el género por defecto, 3) la jerarquía de indicios para asignar un género

sigue este orden: morfosintaxis (concordancia) > morfofonología (terminación del sustantivo) > semántica (género natural).

Si bien los datos presentados hasta aquí ofrecen información detallada sobre las primeras etapas de adquisición de la adquisición del rasgo de género, otros tipos de metodologías orientadas hacia el procesamiento del lenguaje en tiempo real (como las tareas que miden los tiempos de reacción, la técnica de seguimiento ocular o los potenciales relacionados con eventos) nos permiten complementar estas observaciones y también compararlas con el procesamiento del lenguaje en el adulto. Por ejemplo, en un estudio con niños de 3 años (n=26) y adultos (n=26) de español L1, Lew-Williams y Ferrald (2007) utilizaron una tarea de identificación de imágenes mediante la técnica de seguimiento ocular para comparar el procesamiento de la concordancia de género de un grupo a otro. En la tarea, los participantes veían dos imágenes del mismo género o de género diferente y oían una frase que incluía el sustantivo meta precedido por el artículo correspondiente a su género (el/la). El razonamiento de la tarea es que, en los ítems con dos sustantivos del mismo género, el determinante no proporciona ninguna ayuda, pero, en los ítems de diferente género, el determinante facilita la identificación de la respuesta. Los resultados indican que, si bien los adultos identifican la respuesta más rápidamente, ambos grupos sacan provecho del indicio de género en el determinante para identificar la respuesta, ya que el tiempo de reacción es significativamente más corto en la condición de género diferente. Los datos de Jakubowicz y Roulet sobre el proceso de adquisición del francés L1 apuntan en la misma dirección, es decir, hacia la importancia de la concordancia desde muy temprano en el proceso de adquisición: “the fact that 2-year-olds are sensitive to (mis)agreement between the determiner and the noun indicates that the Agree operation is available at a period when determiners are still not consistently produced by the child.” (2007, p. 196). De manera similar, pero esta vez con adultos, el estudio de Faussart et al. (1999) presenta dos experimentos similares, uno con hablantes de español L1 y otro de francés L1. En las dos tareas, las incongruencias de género entre el determinante y el sustantivo –en francés (p. ej., “*la chat*”) y en español (“*la gato*”)– provocan un tiempo de reacción más lento en comparación con los estímulos congruentes (“*le chat*”, “*el gato*”).

A su vez, Caffarra y Barber (2015) investigaron el efecto de las violaciones de concordancia y el efecto de la terminación del sustantivo en el procesamiento mediante el uso de la técnica de potencial relacionado con eventos (*Event-related Potential*, ERP). Para ello, administraron una tarea de juicios de gramaticalidad a 32 adultos hablantes de español L1 y midieron a la vez el

acierto en el juicio gramatical y las respuestas cerebrales. En la tarea, las oraciones correspondían a cuatro condiciones que variaban en función del criterio de transparencia (transparente u opaca) y de concordancia (congruencia o incongruencia entre el determinante y el sustantivo). Con un 94,88 % de aciertos en los juicios de gramaticalidad, quedó demostrado que los participantes conocían el género de los sustantivos utilizados y las respuestas cerebrales permitieron matizar el peso de cada condición sobre el procesamiento. Los participantes mostraron una sensibilidad a la transparencia y a la violación de la concordancia en dos momentos distintos. Concretamente, se observó el patrón LAN-P600 en cuanto al efecto de la concordancia y, además, apareció una señal negativa de mayor amplitud entre 200-500 ms después del inicio con los sustantivos transparentes.⁵¹ Siguiendo el modelo de acceso al rasgo de género propuesto por Gollan y Frost (2001), los autores argumentan que la señal temprana observada en función de la terminación sugiere que el género del sustantivo es recuperado a partir de dos vías paralelas: la vía formal (la terminación) y la vía léxica (el rasgo de género del sustantivo en el léxico mental). En el caso de los sustantivos que tienen una terminación opaca, solo es accesible la vía léxica, mientras que en el caso de los que tienen una terminación transparente, las dos vías se complementan. En cualquier caso, el efecto observado de la terminación no interactúa con el efecto de concordancia, por lo que los autores consideran que la vía de acceso al género del sustantivo no tiene un impacto sobre la concordancia. Si bien se desconoce el papel exacto que juega la terminación en el proceso de acceso al léxico, este estudio muestra que los adultos son sensibles a la morfofonología de los sustantivos en su L1.

2.4.2 En 2L1

En adquisición bilingüe, Maez (1983) analizó la producción de tres niñas bilingües (español-inglés) durante un período de 6 meses (18-24 meses) y observó que estas presentaban sensiblemente las mismas etapas de desarrollo que en los estudios de adquisición monolingüe.⁵² Concretamente, pasaban por una primera etapa de producción de sustantivo sin determinante hacia

⁵¹ El efecto LAN (*Left Anterior Negativity*) se suele asociar con la detección inicial de las violaciones morfosintácticas. Por su parte, el efecto P600 (*positive 600*) se asocia generalmente con el reanálisis y la reparación de los errores sintácticos (Osterhout et al., 2004).

⁵² Durante el período de la investigación, el input en inglés era bastante limitado y procedía principalmente de la televisión; en los tres casos, ambos padres interactuaban casi exclusivamente en español con sus hijas.

los 18 meses y luego, a partir de los 19-20 meses de edad, aparecían los determinantes. El autor observa pocos errores de concordancia de género en sus datos, pero cuando los hay, suelen ser con un determinante masculino y un sustantivo femenino, pero no al revés.

El estudio de Fernández Fuertes et al. (2016) a partir de los datos de producción de Simon y Leo, dos gemelos bilingües español-inglés, sobre un período de 5 años (1;01 a 6;11), muestra patrones similares, pero con un mayor nivel de precisión en el análisis y un rango de edad más amplio.⁵³ En la primera etapa (1;01–2;04), los niños presentan una tasa elevada de omisión de determinantes, omisiones que desaparecen en la segunda etapa (2;05–2;10) para uno de los dos niños y en la tercera etapa (2;11–3;04) para el otro. Con la mayor producción de determinantes, ambos niños producen errores de concordancia en la segunda y la tercera etapa, y, a partir de la cuarta etapa (3;05–4;00), los errores están casi ausentes de la producción de ambos niños, por lo que las autoras concluyen que, a partir de los 3 años de edad, los mecanismos de concordancia de género están ya configurados como en la gramática de un adulto. Sobre los errores de concordancia producidos, que representan apenas el 4,8 % de todos los SD producidos, aparece una clara sobregeneralización del masculino (determinante masculino con un sustantivo femenino) y una tasa de error más elevada con los sustantivos inanimados femeninos. Este efecto de la semántica sobre la concordancia parece ir en contra de lo que Pérez-Pereira (1991) observó con los monolingües, pero es posible que la semántica sea relevante sobre todo al inicio del proceso y el estudio de Pérez-Pereira no lo haya captado, al incluir solamente niños de más de 4 años. En cuanto a la terminación del sustantivo, las autoras observan algunos casos de sobregeneralización de las terminaciones transparentes (-o, -a) a determinantes y sustantivos (p. ej., “uno azoto”, “una serpiente”), pero no notan ningún efecto significativo de la morfofonología sobre la distribución de los errores, y, por lo tanto, concluyen que la morfología no guía la concordancia. Estos resultados de adquisición bilingüe dominantes de español concuerdan a grandes rasgos con los hallazgos de los estudios de adquisición monolingüe del español.

Pasando a los datos experimentales, Cuza y Pérez Tattam (2016) usaron una tarea de denominación de imágenes para elicitar SD con niños bilingües español-inglés de 5-10 años ($n=32$)

⁵³ Estrategia una persona=una lengua (padre en español, madre en inglés). Los datos proceden del corpus FerFuLice en CHILDES (MacWhinney, 2000)

y monolingües de español L1 de la misma edad.⁵⁴ En el experimento, los niños tenían que producir a la vez la concordancia entre el determinante y el sustantivo (“la nube”) y entre el sustantivo y el adjetivo (“la nube roja”). Para concentrarse en el efecto del masculino por defecto y no incluir el efecto de la terminación en su estudio, los autores eligieron exclusivamente sustantivos que tenían una terminación opaca (-e, -z, -d, -l o -n). Además de analizar la concordancia de género, los autores se interesaron por el orden de palabra adjetivo-sustantivo, para ver si el orden del inglés (“*a red cloud*”) tenía un impacto sobre las respuestas de los bilingües. Los resultados de los bilingües muestran una sobregeneralización del determinante masculino con los sustantivos femeninos (aunque menos importante de lo esperado) y, sobre todo, de los adjetivos masculinos con los sustantivos femeninos. Finalmente, en cuanto al orden de palabras, los autores observan una correlación importante entre estos errores y los errores de concordancia (p. ej., “*el rojo nube”), por lo que concluyen que el rasgo de género está estrechamente vinculado con el movimiento del adjetivo.

A su vez, Pérez-Tattam et al. (2019) se interesaron por la asignación del género en palabras inventadas con niños de 5-6 años que correspondían a dos perfiles distintos: bilingües simultáneos euskera-español dominantes de euskera (n=21) y monolingües español L1 (n=28). En el experimento, los niños tenían que producir un SD del tipo “el nolete rojo” (donde “nolete” corresponde a la palabra inventada) a partir de estímulos manipulados en función de las siguientes condiciones: género semántico (sexo masculino o femenino), morfofonológico (terminación -o, -a, -e) y sintáctico (marca de género en el determinante y/o el adjetivo). Todos los participantes mostraron un efecto del masculino por defecto, pero de manera más marcada en el grupo de bilingües (83 % de respuestas en masculino vs. 63 % en el grupo de español L1). El experimento presentaba estímulos que correspondían a cada género en proporciones equivalentes, es decir, 50 % en masculino, 50 % en femenino. En menor medida, pero aún de manera significativa, el género de las respuestas de esos bilingües estaba modulado por la terminación (la -a favorecía el uso del femenino) y el género natural del ítem (los ítems femeninos favorecían el uso del femenino). Los indicios sintácticos (marcas de género en el determinante y/o el adjetivo del estímulo) solo presentaron un efecto significativo sobre el grupo de monolingües y solo marginalmente

⁵⁴ Estos bilingües son considerados como hablantes de herencia, es decir, la lengua hablada en casa es el español, pero, dado que viven en los Estados Unidos, la lengua de enseñanza y la lengua mayoritaria en la sociedad es el inglés.

significativo sobre las respuestas del grupo bilingüe. Este estudio muestra que la sobregeneralización del masculino puede extenderse después de las etapas iniciales de adquisición del género (después de los 3 años) cuando el niño se ve obligado (en un contexto experimental) a asignar un género a un sustantivo nuevo. Este efecto es parecido a lo que Pérez-Pereira (1991) observó en español L1, pero el estudio de Pérez-Tatam et al. (2019) permite contrastar el peso de esta sobregeneralización en un grupo de bilingües dominante en una lengua diferente del español (euskera) frente a un grupo de monolingües.

Caffarra et al. (2017) también llevaron a cabo un estudio con bilingües simultáneos euskera-español, pero se interesaron de manera más específica por la interacción de la lengua dominante con la terminación del sustantivo y la sensibilidad a las violaciones de concordancia de género. Para ello, administraron una tarea de juicios de gramaticalidad a un grupo de adultos bilingües dominantes de euskera (n=24) y a otro de dominantes de español (n=24) y midieron a la vez el acierto en el juicio gramatical y las respuestas cerebrales (ERP).⁵⁵ En la tarea, las oraciones correspondían a cuatro condiciones que variaban en función del criterio de transparencia (transparente u opaca) y de concordancia (incongruencia o congruencia entre el determinante y el sustantivo). También utilizaron una prueba cronometrada en la cual los participantes tenían que elegir el género (masculino o femenino) de los sustantivos utilizados en el experimento anterior, pero, esta vez, fuera de contexto. La segunda tarea mostró que, si bien ambos grupos obtuvieron una tasa alta de acierto en la asignación de género (97 % para el grupo de dominantes de español, 96 % para los dominantes de euskera) y mostraron un efecto significativo de la terminación sobre el acierto y el tiempo de reacción, los bilingües dominantes de euskera tuvieron un tiempo de reacción más lento de manera general en comparación con los bilingües dominantes de español. En cuanto a los juicios de gramaticalidad, ambos grupos obtuvieron una alta tasa de aciertos (96 %), pero, en cuanto a las respuestas cerebrales, se observaron diferencias importantes en cuanto al efecto de la terminación. El grupo de bilingües dominantes de español se comportó de manera comparable a los monolingües del estudio de Caffarra y Barber (2015) que comentamos en la sección anterior, es decir, mostraron el patrón LAN-P600, con una señal negativa temprana (a partir de 200 ms) modulada en función de la transparencia de la terminación, pero sin interacción entre

⁵⁵ Los participantes de ambos grupos tenían un alto dominio de ambas lenguas, pero los autores separaron los grupos en función de su uso de la lengua.

los dos efectos. En cambio, el grupo de bilingües dominantes de euskera mostró una interacción entre los dos efectos. Concretamente, las violaciones de concordancia de género con los sustantivos de terminación opaca no provocaron las mismas respuestas cerebrales que las violaciones de género con un sustantivo de terminación transparente, solo elicitaron el efecto P600, sin el LAN (Caffarra et al., 2017, p. 12). Esta diferencia de procesamiento se podría atribuir, según los autores, a una diferencia en la representación del género en el léxico mental de cada tipo de bilingüe. Los autores proponen que los bilingües dominantes de euskera tienen una representación menos estable del género y recurren a otros mecanismos para compensar. Dado que el euskera es una lengua aglutinante en la cual cada morfema que se añade al final de las palabras tiene su significado y la correspondencia entre la forma y la función de estos morfemas es muy alta, es posible que estos bilingües apliquen al español sus rutinas morfosintácticas del euskera, sobreestimando cada indicio potencial que pueda asociarse con un rasgo (-a~femenino, -o~masculino). Si bien los resultados de este estudio no implican ninguna diferencia en la superficie, ya que ambos tipos de bilingües muestran un dominio nativo de la concordancia de género en sus juicios de gramaticalidad y en la tarea de asignación de género, los resultados de las respuestas cerebrales muestran matices importantes: incluso entre bilingües simultáneos que tienen un alto nivel de dominio de sus dos lenguas, existen diferencias en el procesamiento de los indicios de género en función de su uso de la lengua.

Otros estudios se interesaron por el dominio del género gramatical en función de la edad de adquisición, comparando diferentes perfiles de adultos: monolingües, bilingües simultáneos (hablantes de herencia) y aprendientes de L2. Con el objetivo de explorar la interacción entre la edad de adquisición y factores como el género por defecto, el dominio de la concordancia (determinante o adjetivo) y la terminación (transparente u opaca), Montrul et al. (2008) administraron tres tareas a un grupo de hablantes de inglés L1-español L2 (n=72), otro grupo de hablantes de español como lengua de herencia (n=67) y un grupo de control de hablantes nativos que crecieron como monolingües (n=22). Dos de las tres tareas eran escritas y la tercera era oral, para averiguar también si la modalidad de la tarea tenía un efecto. El primer efecto observado en función de la edad de adquisición tiene que ver con la modalidad: los hablantes de herencia, si bien cometieron más errores en la modalidad escrita, mostraron menor variabilidad entre las dos modalidades, mientras que los hablantes de español L2 mostraron una gran variación, con resultados más bajos en la modalidad oral. En cuanto a los factores intralingüísticos, las autoras

observaron que los dos grupos experimentales cometían más errores de concordancia con el femenino que con el masculino, más errores con los adjetivos que con los determinantes y más errores con los sustantivos de terminación opaca que transparente, pero no observaron una diferencia significativa entre los sustantivos animados o inanimados. La proporción más elevada de errores ocurrió cuando se combinaban los tres factores inhibidores, es decir, con los adjetivos que tenían que concordar con un sustantivo femenino de terminación opaca. El estudio de Alarcón (2011), con un diseño similar al de Montrul et al. (2008), llegó a las mismas conclusiones, con los mismos tipos de errores y la misma ventaja en la modalidad oral para los hablantes de herencia.

Varios estudios sobre los bilingües simultáneos se interesaron por la concordancia de género en el marco de un fenómeno común en las comunidades bilingües: la alternancia de código (*code-switching*). Estos estudios que implican generalmente la producción de sintagmas determinantes mixtos de una lengua con género (español) y una lengua que no lo materializa (inglés, euskera), ofrecen otra perspectiva sobre el manejo del género gramatical en un sistema bilingüe. En los datos espontáneos disponibles sobre los SD mixtos (*switched DPs*) español-inglés, existe una clara preferencia hacia el uso del determinante en español y el sustantivo en inglés, como en (2.16), y no al revés, como en (2.17), y esto tanto en los adultos como en los niños (Liceras, Fernández Fuertes, et al., 2008; Liceras et al., 2016).

- (2.16) a. la *chair* / la *girl*
- b. el *pencil* / el *boy*
- c. el *chair* / el *girl*
- d. la *pencil* / la *boy*
- (2.17) a. *the* silla / *the* chica
- b. *the* lápiz / *the* chico

A la hora de determinar qué influencia el género del determinante elegido, los estudios sobre los SD mixtos se han interesado sobre todo por los siguientes factores: el criterio analógico, el masculino por defecto y la morfofonología. Jake et al. (2002) analizan los datos de un corpus formado por la producción de 10 informantes latino-americanos que viven en los Estados Unidos. En ese corpus, se considera que el criterio analógico explica el 36 % de los SD mixtos en masculino (p. ej., “el *weekend*”, “el fin de semana”) y en femenino (p. ej., “la *broom*”, “la escoba”), la morfofonología explica el 23 % y la combinación de los dos factores explica otro 15 %. Para los 24 % restantes, los autores consideran que los hablantes utilizan el masculino como género por

defecto, ya que solo encuentran una ocurrencia de SD mixto con un determinante femenino (“una rent”, “un alquiler”).⁵⁶ Por su parte, Otheguy y Lapidus Shin (2003) analizan un corpus formado por la producción de 33 informantes latino-americanos que viven en Nueva York y concluyen que el masculino por defecto es el factor más determinante, ya que el 87 % de los SD mixtos que cuentan con un determinante son masculinos, una tasa igual a la obtenida por Poplack et al. (1982) en un corpus similar. Los pocos casos de sustantivos que reciben el género femenino se deben a la terminación en -a (p. ej. “la *boila*”, una adaptación del inglés “boiler”) o al género natural del referente (p. ej., “una *teenager*”, en referencia a una adolescente). Sin embargo, incluso el género natural del referente es a veces desplazado por el masculino por defecto en algunos casos, como “un *social worker*” o “los *mid-wives*”. En cuanto al criterio analógico, los autores encuentran tres ejemplos que solo se podrían atribuir a ese factor: “la *liquor store*” y “una *deli*” por la posible analogía con “la tienda” y “la *high school*” por la posible analogía con “la escuela”. En cambio, según el análisis de Liceras et al. (2006) de los datos de Moyer (1993) sobre el español de Gibraltar, el género analógico explica el 63 % de la asignación de género en los SD mixtos del corpus ($n=206$). Todos estos datos espontáneos sobre los SD mixtos pueden parecer contradictorios, pero muestran que las estrategias de asignación del género en los SD mixtos pueden variar de un contexto lingüístico a otro.

Los resultados experimentales de Liceras et al. (2012) y de Valenzuela et al. (2012) reportados en Liceras et al. (2016) muestran que, si bien los niños bilingües simultáneos inglés-español (dominantes de inglés) presentan cierta sensibilidad al criterio analógico, tienden a aceptar a menudo el masculino por defecto, una preferencia distinta de los bilingües secuenciales dominantes de español, como veremos en la próxima sección sobre la L2.⁵⁷ Esta preferencia por el masculino por defecto parece aún más marcada en los bilingües simultáneos llegados a la edad adulta (Valenzuela et al., 2012).

⁵⁶ Consideramos que esta excepción al masculino por defecto se podría atribuir a la semejanza entre la forma de “rent” en inglés y “renta” en español.

⁵⁷ Además de la concordancia (*concord*) en los SD mixtos con un determinante en español y un sustantivo en inglés (p. ej., el *book*), Liceras et al. (2016), como en varios de sus estudios experimentales anteriores, se interesan también por los SD mixtos con el determinante en inglés (p. ej., *the libro*) y por la concordancia (*agreement*) entre un SD en una lengua y un predicado adjetivo en la otra (p. ej., *the book es bonito*).

Al contrario de la tendencia del masculino por defecto observada en los SD mixtos de los bilingües inglés/español, Parafita Couto et al. (2016) observó que, en general, los bilingües español-euskera (n=30, de los cuales 26 son dominantes de español) que participaron en su estudio preferían el determinante femenino en las estructuras que combinaban un determinante en español y un sustantivo en euskera (p. ej. la *tipula* “cebolla”, la *gezia* “flecha”). Los autores proponen que dicha preferencia podría proceder del hecho que la terminación -a, que es frecuente en euskera porque puede representar a la vez la terminación del sustantivo (-a léxica) o el artículo definido (el, la).⁵⁸ El experimento explota esta particularidad de la lengua, con estímulos con la -a léxica (el/la *tipula*), pero también con otros que presentan artículos duplicados donde aparece a la vez el determinante del español y el determinante del euskera (el/la *gezia*). Además de lo que los autores consideran como el femenino por defecto, el género analógico también tiene un efecto significativo sobre la preferencia de los participantes.

Badiola y Sande (2018) también se interesaron por la preferencia de los bilingües español/euskera (n=21) en una tarea de juicios de aceptabilidad con un SD mixtos (determinante en español, sustantivo en euskera).⁵⁹ Los participantes mostraron una preferencia por el determinante masculino de manera general y una preferencia por el determinante femenino solo cuando el sustantivo en euskera terminaba por una -a léxica (posiblemente por analogía con la terminación -a del español). El criterio analógico parece tener una importancia menor. Las autoras proponen dos posibles razones para explicar la diferencia entre sus resultados y los de Parafita Couto et al. (2016). Por una parte, los participantes de Parafita Couto et al. (2016) parecen interpretar el doble determinante (-a no léxica) como si fuera una -a léxica y, por lo tanto, aceptan tanto “la *tipula*” como “la *gezia*”, lo que desequilibra los datos a favor de la terminación en -a (la casi totalidad de los ítems aceptados terminan en -a). En cambio, los participantes del estudio de Badiola y Sande (2018) tienden a rechazar los SD con el doble artículo, por lo que los ítems

⁵⁸ Por ejemplo, en “*tipula*” (cebolla) la -a es léxica, forma parte del sustantivo (**tipul*), pero en “*gezia*” (la flecha), la -a representa el artículo porque el sustantivo solo es “*gezi*”. Cuando un sustantivo termina en -a léxica, no se duplica la -a cuando se añade el artículo, por lo que “*tipula*” se puede interpretar como “cebolla” o “la cebolla”, según el contexto.

⁵⁹ Los resultados de los participantes en las pruebas de niveles en euskera y en español revelan que su perfil de bilingüismo parece equilibrado. Sin embargo, los autores no analizaron el efecto de la lengua hablada en casa por los participantes sobre sus resultados, pero por lo menos reportan que 14 de ellos hablan español y euskera, 2 solo euskera y 5 solo español.

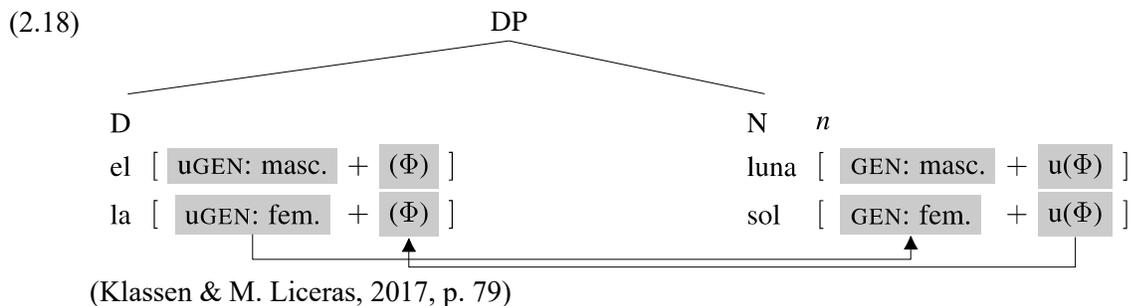
aceptados están más equilibrados entre la -a léxica y las demás terminaciones. Por otra parte, las autoras observan que los bilingües de cada estudio corresponden a un perfil diferente. Los participantes de Parafita Couto et al. (2016) son principalmente dominantes de español, con una edad de adquisición del euskera en torno a los 3 años, en contexto escolar, lo que hace de ellos unos bilingües tempranos secuenciales y no simultáneos. En cambio, los bilingües del estudio de Badiola y Sande (2018) son bilingües simultáneos que crecieron en un entorno vasco parlante (Guernica). Esta diferencia en el desarrollo de las dos lenguas de los bilingües parece conllevar diferencias profundas en el tratamiento de lo que es, en realidad, la L2 de unos hablantes versus la 2L1 de otros. A la luz de estos resultados, lo que Parafita Couto et al. (2016) observaron no parece ser un “femenino por defecto” sino una simple extensión de las reglas morfofonológicas del español (la lengua dominante de los hablantes) a palabras de su L2.

Para completar esta sección, presentamos datos sobre la alternancia de código en los bilingües secuenciales que, si bien entrarían mejor en la sección siguiente sobre la L2, nos permitirán hacer un balance sobre los estudios de alternancia de código antes de pasar a trabajos basados en otras metodologías.

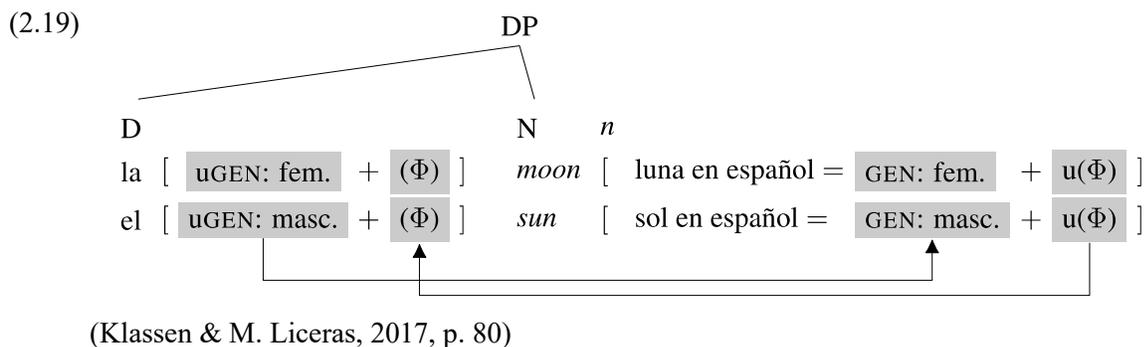
En lo referente a las diferentes estrategias utilizadas por los bilingües secuenciales y simultáneos a la hora de asignar un género a los SD mixtos en un contexto experimental, Liceras et al. (2016) presentan un resumen de las diferencias observadas en función del criterio analógico y el masculino por defecto. Como vimos en los estudios citados anteriormente, los bilingües simultáneos inglés-español (dominantes de inglés) muestran una menor sensibilidad al criterio analógico y una mayor preferencia por el masculino por defecto. En cambio, los bilingües secuenciales, tanto niños (Liceras et al., 2012) como adultos (Liceras, 2013; Liceras, Zobl, et al., 2008; Valenzuela et al., 2012), muestran una clara preferencia por el criterio analógico cuando son dominantes de español, pero los resultados son menos claros cuando son dominantes de inglés (Klassen et al., 2014; Liceras, Fernández Fuertes, et al., 2008), ya que parecen usar a la vez el criterio analógico y el masculino por defecto. El estudio experimental de Klassen y Liceras (2017) parece confirmar esta tendencia observada en función de la lengua dominante. Los bilingües español L1/inglés L2 (n=43) muestran una clara preferencia por el criterio analógico, mientras que los bilingües de inglés L1/español L2 (n=18 de nivel intermedio, n=25 de nivel avanzado) no muestran variación significativa en función del criterio analógico y, si bien una parte de éstos (L2

más avanzados) aceptan más los SD mixtos con un determinante masculino, el efecto del masculino por defecto tampoco es significativo en estos hablantes.

Liceras et al. (2016) argumentan que estas diferencias entre los distintos tipos de bilingües se pueden explicar gracias al Mecanismo de cotejamiento de los rasgos dobles de género (*Gender Double-Feature Valuation Mechanism*; Liceras, Fernández Fuertes, et al., 2008). Según esta propuesta, para establecer la concordancia en español, dos rasgos tienen que ser cotejados: el rasgo de género inherente al sustantivo y el rasgo funcional no interpretable de concordancia de género (Φ) que se ubica en el determinante. El ejemplo (2.18) ilustra el cotejamiento de los rasgos en un SD en español.

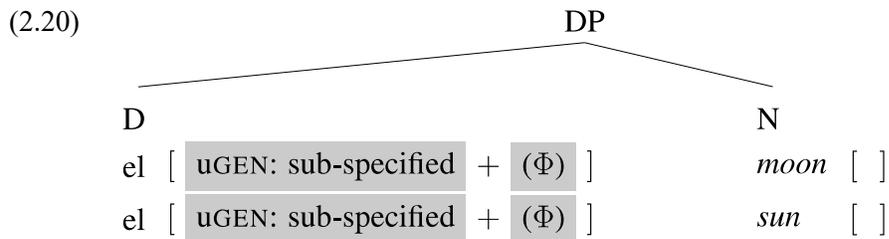


En los SD mixtos, existen dos posibilidades. En los bilingües dominantes de español, que prefieren el criterio analógico, los dos rasgos son cotejados exitosamente, ya que el sustantivo del inglés está asociado al rasgo de género del sustantivo equivalente en español como en el ejemplo (2.19).



En cambio, en los bilingües dominantes de inglés, el sustantivo en inglés no está asociado con ningún género, sino que está subespecificado en cuanto al género en comparación con los sustantivos del español, por lo que, si aparece con un determinante en español, este también tiene

que ser subespecificado y, como la forma por defecto en español corresponde con el masculino, el determinante adopta esta forma, como podemos observar en el ejemplo (2.20).



(Klassen & M. Liceras, 2017, p. 81)

Este razonamiento, combinado con las observaciones sobre el procesamiento del género en Caffarra et al. (2017), también puede ser extendido a los resultados sobre el euskera de Badiola y Sande (2018) y Parafita Couto et al. (2016). La principal diferencia entre la combinación inglés-español y euskera-español parece ser que, en el caso del bilingüismo euskera-español, la morfofonología ejerce un papel más importante que en el caso del bilingüismo inglés-español. Esto se podría atribuir a la mayor frecuencia de la terminación -a en euskera en comparación con el inglés y el hecho de que la terminación -a del euskera coincide en la forma con la terminación típicamente femenina del español. Como los bilingües dominantes de inglés, los dominantes de euskera no parecen mostrar una preferencia por el criterio analógico y recurren, en muchos casos, al masculino por defecto. Sin embargo, en algunos casos, usan el determinante femenino porque asocian al sustantivo del euskera el género femenino, basándose en la forma (la terminación -a). Entonces, estos bilingües prefieren la forma subespecificada del determinante porque, en su léxico, los sustantivos del euskera no suelen estar asociados con un género, pero cuando el sustantivo termina en -a, prefieren usar un determinante especificado con el género femenino, por analogía con la forma del español. En cambio, los dominantes de español, prefieren usar una forma especificada del determinante (masculino o femenino) en todos los casos, por lo que buscan una forma que concuerde en género con el sustantivo. Para asignar el género al sustantivo del euskera, esos bilingües tienen dos estrategias: o bien usan el criterio analógico (como los bilingües español-inglés dominantes de español), o bien reanalizan el sustantivo según los patrones morfofonológicos del español y le asignan el género femenino a las unidades léxicas que terminan en -a. Esta segunda estrategia está poco documentada en el caso de los SD mixtos español-inglés, muy probablemente porque hay pocos sustantivos que terminan en -a en inglés, pero recordamos el ejemplo de OtheGuy

y Lapidus Shin (2003) que comentamos antes, *boila* (adaptación de “boiler”) que aparecía con el género femenino.

2.4.3 En L2

Además de los estudios sobre la alternancia de código, varios estudios se interesaron por la concordancia y el procesamiento del género en español L2 en distintos niveles de dominio de la lengua, inspirándose de los criterios que parecen facilitar o inhibir la adquisición y el procesamiento del género en L1/2L1: el género por defecto, la morfofonología, la concordancia, el criterio analógico y el hecho de hablar una L1 que materializa el género. Con respecto al criterio analógico, si bien este concepto procede de los estudios de alternancia de código con SD mixtos, también se puede aplicar a los SD producidos exclusivamente en la L2 cuando la L1 del aprendiente presenta un sistema de género similar al de la L2.⁶⁰ En cuanto a la importancia de la configuración de género de la L1 de manera más general, este factor es de particular interés cuando el diseño de un estudio no permite evaluar de manera precisa el peso del criterio analógico. A continuación, revisaremos algunos estudios que se basan en datos espontáneos y otros de corte experimental que exploran de manera más específica el papel de la L1. Dado que el papel de la L1 puede ser de gran interés para guiar las hipótesis de influencia hacia la L3 en el caso de nuestro estudio, nos detendremos más en estos estudios.

En un estudio de caso, Andersen (1984) se interesó por la concordancia de género en la producción de un hablante de inglés L1 de 12 años que emigró a Puerto Rico y que adquirió su español L2 en un contexto informal. Los datos muestran una sobreutilización del masculino aplicada a la casi totalidad los modificadores de los sustantivos.

A su vez, Finneman (1992) se interesó por la concordancia de género por parte de tres aprendientes universitarios de nivel inicial de español L2 (inglés L1) a partir de los datos de una serie de entrevistas realizadas sobre un período de seis meses. El autor observó que los aprendientes

⁶⁰ Siguiendo este razonamiento, en esta tesis emplearemos el término “criterio analógico” para referirnos al proceso de asignación del género de una unidad léxica en la L2 a su equivalente en la L1. Como hemos mencionado, es un término que se asocia a los trabajos de alternancia de código y no es de uso común en trabajos de L2/L3. A pesar de ello, consideramos que se trata, en esencia, del mismo proceso y, por lo tanto, vemos adecuado el empleo de dicho término en estos contextos de adquisición.

sobreutilizaban el género masculino, pero también se guiaban por la terminación y el género natural del sustantivo cuando era posible. Según esta observación, el masculino parece ser la opción por defecto cuando no hay otros indicios disponibles. Sin embargo, uno de los participantes muestra una mayor tendencia al sobreuso del masculino, sin mostrar sensibilidad a los demás indicios, tendencia que se acentúa a lo largo del período cubierto por las entrevistas, por lo que el autor considera que el masculino por defecto puede ser una estrategia individual del aprendiente. En cuanto al dominio de la concordancia, se producen menos errores en la concordancia con los determinantes que con los adjetivos.⁶¹

En otro estudio descriptivo de la producción de siete aprendientes de español L2 (inglés L1) de nivel intermedio, Fernández-García (1999) observa que los aprendientes producen más sustantivos con terminación transparente (-o, -a) y que cometen menos errores de concordancia con esos sustantivos y con los sustantivos que tienen un género natural. De manera general, los participantes sobreutilizan los determinantes masculinos, pero algunos de ellos muestran una tendencia a la sobreutilización del femenino. Además, la mayoría de sus participantes cometen menos errores en la concordancia con los determinantes que con los adjetivos.

Bruhn de Garavito y White (2002) exploraron la concordancia de género en español L2 por parte de hablantes de francés L1 (edad: 14-19 años) de dos niveles distintos (inicial=30, intermedio=12) mediante una tarea de producción elicitada a partir de imágenes. El análisis de los SD mostró que 1) los aprendientes podían producir los SD con el orden de palabras adecuado (el mismo que en francés), 2) tenían más dificultad con los artículos indefinidos que con los definidos y aún más con los adjetivos, 3) tendían a sobregeneralizar el uso del masculino en los determinantes y, de manera aún más marcada, en los adjetivos.⁶² Además, de manera un poco contraintuitiva, como comentan las autoras, el grupo de nivel inicial cometió más errores de concordancia con los sustantivos que tenían un género natural (los animados). Esto parece indicar, según las autoras, que el hecho de tener el rasgo abstracto de género en su L1 y de conocer el género del sustantivo (género natural) no siempre es suficiente para realizar la concordancia. Las autoras sugieren que el

⁶¹ En los determinantes, los artículos definidos tienen una tasa de error más baja que los indefinidos. En cuanto a los adjetivos, los que aparecen dentro del SD tienen una tasa de error más baja que los que aparecen en el predicado del verbo.

⁶² Las autoras no reportan los efectos de la transparencia de la terminación y del criterio analógico.

problema de los aprendientes no es un problema ni con el rasgo abstracto de género, ni con el género del sustantivo, sino de la relación (*mapping problem*) entre el rasgo abstracto con la forma superficial del elemento que tiene que concordar. Esta propuesta de Lardiere (2000) sobre la adquisición de L2 está en el centro de la hipótesis de la flexión superficial ausente (*Missing Surface Inflection Hypothesis*; Haznedar & Schwartz, 1997; Prévost & White, 2000). En el caso de los resultados de Bruhn de Garavito y White (2002), el problema de relación entre el rasgo abstracto y la forma sería responsable del sobreuso del masculino, es decir, los errores no se deberían a la ausencia del rasgo abstracto de género, ni a la asignación del género en el léxico, sino a la imposibilidad por el hablante de relacionar, en tiempo real, el rasgo femenino con la forma femenina, lo que resulta en la producción de la forma por defecto, el masculino.

White et al. (2004) se interesaron por la concordancia de género en hablantes de español como L2 comparándolos según su L1 (francés=48, inglés=68) y su nivel de dominio del español (L1 inglés, inicial=34, intermedio=24, avanzado=10, L1 francés, inicial=12, intermedio=20, avanzado=16).⁶³ El objetivo del estudio consistía en averiguar si los hablantes de una L1 sin rasgo funcional de género gramatical (inglés) eran capaces de adquirir los rasgos abstractos de género (interpretable y no interpretable) en una L2 con género gramatical (español) y alcanzar el nivel del grupo de control de hablantes de español L1 (n=20) (de aquí en adelante, nos referiremos a las lenguas como el inglés mediante el rasgo [-gen] y la lenguas como el español y el francés con el rasgo [+gen]). Según la hipótesis de los rasgos funcionales fallidos (*Failed Functional Features Hypothesis*; Hawkins & Chan, 1997), después de la pubertad, resulta imposible adquirir nuevos rasgos abstractos, como el género gramatical, si estos no existen en la L1. En cambio, según la hipótesis de la transferencia completa y del acceso completo (*Full Transfer Full Access Hypothesis*; Schwartz & Sprouse, 1996), en el caso de que la transferencia de los rasgos funcionales a partir de la L1 no resulte exitosa, los rasgos se pueden adquirir posteriormente, sin limitación de edad o de maduración. En las cuatro tareas del estudio (dos de producción, una de vocabulario y una de comprensión), las autoras no observaron una diferencia clara entre los grupos experimentales en función de su L1. Ambos grupos produjeron más errores de concordancia con los adjetivos (tareas de producción), menos errores cuando la terminación del sustantivo era transparente (tarea de

⁶³ Como las autoras comentan, para varios participantes el español no es la L2 sino la L3, por eso, volveremos brevemente sobre este estudio en la sección sobre L3.

vocabulario) y se observó una sobregeneralización del masculino (todas las tareas), sobre todo en la concordancia con los adjetivos. En la tarea de vocabulario, se observó un efecto del criterio analógico con el grupo de francés L1, es decir, este grupo produjo más errores cuando el género del sustantivo en español difería del género del sustantivo equivalente en francés. El criterio analógico afectó de manera significativa a los grupos de francés L1 de nivel inicial e intermedio y, aun en el nivel avanzado, la diferencia de acierto entre los sustantivos con género opuesto (90 %) y el género analógico (99 %) era notable. Considerando las tasas de acierto altas de los grupos de nivel avanzado tanto de inglés L1 como de francés L1 en todas las tareas y la falta de diferencia significativa entre esos grupos y el grupo de control de español L1, las autoras concluyen que el hecho de tener una L1 [-gen] no impide la adquisición del rasgo funcional de género en la L2 [+gen], en contra de la hipótesis de los rasgos funcionales fallidos de Hawkins y Chan (1997). Los resultados tampoco son plenamente compatibles con la hipótesis de la transferencia completa y del acceso completo (Schwartz & Sprouse, 1996), ya que una transferencia inicial debería favorecer al grupo de francés L1, pero no se observa una diferencia clara entre los grupos de nivel inicial en función de su L1, salvo en lo que concierne al criterio analógico. Los resultados apuntan, sin embargo, hacia un acceso completo al inventario de rasgos funcionales, incluso los no disponibles en la L1.

A su vez, Franceschina (2005) se interesó por la concordancia de género en los hablantes de español L2 de nivel avanzado, comparando hablantes de diferentes L1 [+gen] (n=53) con otros hablantes de una L1 [-gen] (n=15).⁶⁴ En el marco de este estudio, que implica varias tareas (espontáneas y experimentales; orales y escritas) y varias combinaciones de lenguas, resulta difícil evaluar la importancia del criterio analógico, ya que un sustantivo dado (p. ej., “maleta”, femenino), puede tener un género en una L1 (“*koffer*”, masculino en alemán) y otro género en otra L1 (“*valise*”, femenino en francés). Sin embargo, la autora se interesó de manera más general por cómo puede contribuir el hecho de materializar el género en la L1 en el dominio de la concordancia de género en la L2. Concretamente, se observó que el grupo de L1 [+gen] tenía un dominio comparable al de los nativos en todas las tareas, mientras que el grupo de L1 [-gen] cometía más errores y esos errores se veían afectados de manera significativa por la morfofonología, es decir, la transparencia

⁶⁴ Las diferentes L1 [+gen] eran el árabe, el francés, el alemán, el griego, el italiano y el portugués y la única L1 [-gen] era el inglés. La autora constató estos datos con un grupo de control formado por hablantes de español L1 (n=42)

o la opacidad de la terminación del sustantivo. Sobre la base de sus resultados, la autora concluye que los hablantes de español L2 pueden alcanzar un nivel casi nativo en cuanto al género gramatical si tienen una L1 [+gen], pero que este nivel es de más difícil alcance para los hablantes de una L1 [-gen]. Además, el impacto significativo de la terminación sobre solo el grupo de L1 [-gen] parece indicar que el hecho de tener una L1 [-gen] o [+gen] condiciona las estrategias privilegiadas a la hora de asignar o procesar el género en la L2.

De manera similar, Sabourin et al. (2006) se interesaron por la concordancia de género, pero esta vez con hablantes de neerlandés L2 de nivel avanzado con diferentes L1 [+gen] que tienen un sistema similar (masculino, femenino, neutro; alemán; n=25) o diferente del neerlandés (masculino/femenino; español, francés o italiano; n=21), y otros hablantes de una L1 [-gen] (inglés; n=24).⁶⁵ A partir de una tarea de asignación en la cual los participantes solo tenían que elegir el determinante del género adecuado y otra tarea de concordancia a larga distancia con un pronombre relativo, los autores exploraron los efectos de la L1 en función del rasgo [\pm gen], del género por defecto, de la frecuencia de los sustantivos y del contraste entre asignación y concordancia. En cuanto al género por defecto, todos los grupos muestran una tendencia a sobreutilizar el género común, particularmente con los sustantivos menos frecuentes. Con respecto a la diferencia entre asignación y concordancia, el efecto es parecido al contraste entre la concordancia con el determinante y el adjetivo observado en estudios sobre el español, es decir, las tasas de acierto son más elevadas en la asignación. En cuanto al efecto de la L1, en la tarea de asignación, se observó un efecto facilitador en el caso del grupo de alemán L1 y, en menor medida, en el caso de las demás L1 [+gen]. En la tarea de concordancia, la diferencia en función de la L1 [+gen] (alemán vs. español, francés e italiano) disminuyó. En ambas tareas, el grupo de inglés L1 obtuvo los resultados más bajos. Las autoras concluyen entonces que la L1 puede influenciar de dos maneras distintas: con una transferencia profunda cuando las lenguas tienen el rasgo abstracto de género gramatical y de manera superficial cuando las categorías de género son las mismas en las dos lenguas. El grupo de alemán se beneficia del efecto facilitador de la influencia en los dos niveles, mientras que

⁶⁵ Si bien el sistema de género del neerlandés ha evolucionado hacia un sistema dual con 1) el género común en el cual los géneros masculino y femenino han colapsado y 2) el neutro, sigue compartiendo mucho en común con el sistema de tripartita del alemán. Por ejemplo, la mayoría de los sustantivos de género masculino o femenino del alemán corresponden con el género común en neerlandés, los patrones de asignación y de concordancia son parecidos (Sabourin et al., 2006, p. 6)

el grupo de lenguas romances L1 solo se beneficia de la transferencia en el nivel profundo y el grupo de inglés L1 no se beneficia de ningún efecto facilitador. Estos resultados son compatibles con los de Franceschina (2005), que observó un efecto facilitador de la L1 [+gen] sobre la L2 [+gen], pero van en contra de White et al. (2004), que no observaron dicho efecto facilitador en función de la L1 [\pm gen]. Consideramos que la ausencia de diferencias significativas en el caso de White et al. (2004) podría estar relacionada con el nivel de dificultad de la tarea, es decir, que el efecto techo (resultados cerca del 100 %) observado en los grupos de nivel avanzado no permite diferenciar de manera significativa al grupo de inglés L1 del grupo de francés L1.⁶⁶ En cambio, en los estudios de Franceschina (2005) y Sabourin et al. (2006), las tareas permitieron observar mayor variación en los resultados, gracias a lo cual se pudo observar un efecto de la L1.

Para ahondar en las posibles diferencias entre hablantes de L2 de nivel casi nativo y hablantes nativos, los estudios sobre el procesamiento ofrecen una perspectiva que permite resaltar diferencias a veces imperceptibles en las tareas *offline*. Por ejemplo, Keating (2009) utilizó la técnica de seguimiento ocular para estudiar la sensibilidad a las violaciones de concordancia de género con adultos hablantes de inglés L1–español L2 (niveles inicial, n=18; intermedio; n=14; avanzado, n=12) y hablantes nativos de español (n=18). La tarea consistía en leer oraciones con errores de concordancia en los adjetivos dentro del SD (p. ej., “una casa *blanco”), en el predicado del verbo (p. ej., “una casa es bastante *pequeño cuando...”) o en una cláusula de relativo (p. ej., “una casa se vende bastante bien cuando es *nuevo y muy grande”). Los grupos de nivel inicial no mostraron una sensibilidad significativa a la violación de concordancia ni dentro, ni fuera del SD. En cambio, el grupo de nivel avanzado mostró una sensibilidad parecida al grupo nativo frente a las violaciones de concordancia dentro del SD, pero diferente del grupo nativo cuando la violación ocurría fuera del SD. El autor concluye que los hablantes de L2 de nivel avanzado logran adquirir el rasgo abstracto de género, ya que tienen un dominio nativo de la concordancia dentro del SD, lo cual es incompatible con la hipótesis de los rasgos funcionales fallidos (Hawkins & Chan, 1997) y coincide con los resultados de White et al. (2004). Keating subraya una diferencia entre sus hablantes de nivel avanzado y los de estudios anteriores (Franceschina, 2001, 2005;

⁶⁶ Montrul et al. (2008) comentan al respecto del nivel de dificultad de la tarea de producción oral de White et al. (2004) que es muy probable que, en un contexto de producción libre, los hablantes hayan producido principalmente sustantivos muy frecuentes con terminación transparente, lo que favoreció una tasa de acierto más elevada.

Sabourin et al., 2006) que mostraron una diferencia entre los nativos y los hablantes de nivel avanzado: la experiencia del aprendiente. Todos los participantes de nivel avanzado del estudio de Keating (2009) habían completado al menos un grado universitario en español, habían vivido y/o estudiado en un país de habla hispana y la mayoría de ellos estudiaban una carrera académica en español en el momento del estudio. Por lo tanto, su experiencia con la lengua española incluía, entre otras cosas, varias lecturas para fines académicos, redacción de ensayos e interacciones frecuentes con hablantes nativos, lo cual, según Keating, les permitió adquirir un nivel casi nativo. Si bien este nivel casi nativo se caracteriza por la adquisición de rasgos funcionales abstractos que permiten procesar el género dentro del SD como los nativos, el grupo de nivel avanzado todavía muestra diferencias en el procesamiento de la concordancia a larga distancia, fuera del SD. Esta diferencia se puede explicar mediante la hipótesis de la estructura superficial (*Shallow Structure Hypothesis*; Clahsen & Felser, 2006a, 2006b), es decir, un déficit de procesamiento que impide el análisis de las relaciones a larga distancia en la L2, posiblemente causado por la carga cognitiva adicional que representa el procesamiento en la L2.

Gillon Dowens et al. (2009) también exploraron la sensibilidad a las violaciones de concordancia de género con adultos hablantes de inglés L1–español L2 de nivel avanzado (n=33) en comparación con un grupo de español L1 (n=33), pero esta vez observando las respuestas cerebrales (ERP) en el marco de una tarea de juicios de gramaticalidad. Los participantes tenían que juzgar oraciones en las que las violaciones de género podían aparecer o al principio (“* la suelo está plano y bien acabado”) o en el medio (“el suelo está *plana y bien acabado”) de la oración. El grupo nativo mostró sensibilidad a las violaciones en las dos posiciones, mientras que el grupo de español L2 mostró un efecto diferente en las dos posiciones, con una sensibilidad similar al grupo de nativo en la primera posición, pero distinta en la segunda. Los autores argumentan que este efecto no se debe a la ausencia del rasgo abstracto de género en la gramática del hablante, sino a un efecto de procesamiento en la L2 relacionado con otros aspectos cognitivos, como la memoria de trabajo.

La cuestión del procesamiento es un tema recurrente en la literatura sobre la concordancia de género que compara hablantes nativos y no nativos. Algunos de los estudios que revisamos en la presente sección sugieren que las diferencias que presentan los no nativos se deben a un efecto de procesamiento y no tanto a una ausencia del rasgo abstracto de género en la gramática del hablante no nativo, al menos en un nivel avanzado (Gillon Dowens et al., 2009; Keating, 2009;

White et al., 2004), mientras que otros sugieren que los hablantes de una L1 [-gen] pueden difícilmente adquirir los rasgos de género de una L2 [+gen] (Franceschina, 2005; Sabourin et al., 2006). En cuanto a la influencia de la L1, varios estudios sugieren un efecto facilitador cuando la L1 y la L2 tienen sistemas de género gramatical similares (Franceschina, 2005; Sabourin et al., 2006), mientras que otros no observan este efecto facilitador (White et al., 2004). Sin embargo, cabe mencionar que aun en el caso de White et al. (2004) se observa un efecto del criterio analógico, por lo que la influencia de la L1, incluso si no opera sobre el proceso de adquisición del rasgo funcional de género, al menos parece tener una influencia a nivel léxico. En cuanto a los factores intralingüísticos, la mayoría de los estudios sugieren que la morfofonología, la concordancia y, en menor medida, la semántica parecen funcionar como indicios que facilitan a la vez la adquisición y el procesamiento del género en español. Ante la ausencia de indicios suficientes, o en casos en los que no se puede procesar toda la información en tiempo real (concordancia con los adjetivos en las etapas iniciales, concordancia de larga distancia hasta el nivel avanzado), los hablantes de L2 sobreutilizan el género masculino, ya que representa el género por defecto en español. La adquisición de L2 se guía esencialmente por los mismos indicios que la adquisición de la L1, pero el resultado alcanzado es distinto. El efecto de los factores intralingüísticos sobre el procesamiento de la L1 desaparece prácticamente después de las primeras etapas de adquisición, pero dado que el procesamiento en L2 representa, de manera general, un coste cognitivo adicional (McDonald, 2006), esos efectos perduran en la lengua no nativa, hasta niveles avanzados.

2.4.4 En L3

Hasta donde sabemos, no existen estudios sobre la adquisición de español L3 que se hayan interesado por la concordancia de género dentro del SD con una combinación de lenguas como la que tenemos en nuestro estudio, es decir, una L1/L2 que materializa el género y otro que no lo materializa. Por lo tanto, solo revisitamos los resultados de White et al. (2004) que hemos comentado en la sección anterior, pero ahora con la perspectiva de la L3 y añadimos un detalle sobre un estudio citado por Liceras et al. (2016) en cuanto a la alternancia de código.

En White et al. (2004), las autoras mencionan que el español es la L3 de ciertos participantes de inglés L1, ya que han adquirido el francés como L2. Sin embargo, los resultados no revelan ningún efecto aparente del criterio analógico (género del francés vs. el español) sobre estos participantes, mientras que este sí afecta de manera significativa al grupo de francés L1. Por su

parte, Liceras et al. (2016) comentan un estudio de Llama et al. (2011) sobre la alternancia de código con hablantes de inglés L1-francés L2-español L3 y mencionan que estos participantes, en sus SD mixtos, no tendían a usar el criterio analógico sino el masculino por defecto. Esto parece indicar que, cuando la lengua [+gen] no es la L1 sino la L2, el criterio analógico no se aplica, como tampoco parece ocurrir cuando el español no es la lengua dominante de los bilingües español-inglés.

2.5 Síntesis provisional

Como hemos visto hasta aquí, el género gramatical es un fenómeno complejo ampliamente estudiado tanto desde un punto de vista teórico (secciones 2.2 y 2.3) como desde el campo de la adquisición (sección 2.4).

Nuestro trabajo sigue la línea trazada por los estudios de adquisición de L1, 2L1 y L2 sobre el tema y busca ampliarla añadiendo combinaciones de lenguas no estudiadas hasta la fecha, como es el caso del bilingüismo euskera-francés. La revisión de los trabajos anteriores nos permitió destacar que, en lo referente al género gramatical, existe una tendencia a sobreutilizar el género masculino en la concordancia en ciertos contextos precisos, nos referimos a esto como **el género (masculino) por defecto**. En L1, puede ocurrir en las primeras etapas de adquisición de la lengua (Schnell de Acedo, 1994) o cuando se tiene que hacer concordar una nueva palabra cuyo género es desconocido, como un préstamo o una palabra inventada (Pérez-Pereira, 1991). En casos de bilingüismo, esta estrategia se da a menudo en situaciones de alternancia de código, sobre todo cuando la lengua dominante no es el español (Liceras, Fernández Fuertes, et al., 2008; Valenzuela et al., 2012). También se da en situaciones en las que el bilingüe no logra procesar toda la información de manera adecuada, ya sea por cuestiones generales de procesamiento de la L2 (Keating, 2009) o por cuestiones de distancia entre el sustantivo y el elemento que tiene que concordar (Gillon Dowens et al., 2009). Otra observación recurrente en los trabajos de adquisición concierne al **tipo de concordancia**, es decir, la diferencia entre la concordancia con el determinante, con un adjetivo dentro del SD o con un adjetivo u otros elementos fuera del SD. En L1, estas diferencias se observan en las primeras etapas de adquisición y parecen indicar que la concordancia con el determinante tiene un vínculo estrecho con la asignación de género y, por lo tanto, ocurre antes de la concordancia con el adjetivo (Hernández Pina, 1984). En L2, la concordancia con el adjetivo suele presentar más dificultades, sobre todo en los niveles iniciales

(Finnemann, 1992), pero también en los niveles avanzados cuando se observa el procesamiento mediante técnicas *online* (Gillon Dowens et al., 2009). El papel de la **morfofonología**, es decir, la **transparencia de la terminación** del sustantivo también es un fenómeno ampliamente documentado en los trabajos revisados. Si bien su papel en la L1 y la 2L1 no está claramente definido (Fernández Fuertes et al., 2016; Pérez-Pereira, 1991), ciertos estudios sobre el procesamiento muestran que los hablantes de español L1 son sensibles a la diferencia entre una terminación opaca y transparente (Caffarra & Barber, 2015). En L2, varios estudios muestran que los aprendientes de L2 también son sensibles a este factor de la morfofonología (Fernández-García, 1999; Finnemann, 1992; Franceschina, 2005). Finalmente, otro elemento que sobresale de los estudios sobre la alternancia de código y de ciertos estudios de L2 es el **criterio analógico**. En alternancia de código, los bilingües dominantes de español muestran una preferencia por el género analógico del equivalente en español cuando usan un sustantivo de otra lengua con un determinante del español en SD mixtos (p. ej., la *chair*) (Liceras et al., 2016; Liceras, Fernández Fuertes, et al., 2008). En los estudios sobre el español como L2, se alude al criterio analógico cuando el género de un sustantivo en español corresponde al género del sustantivo equivalente en la L1 del hablante (White et al., 2004). Algunos de los trabajos revisados también abordaron el tema de una **influencia de la L1** de manera global cuando se consideraba que los rasgos funcionales de género presentes en la L1 favorecían la adquisición de los rasgos equivalentes en la L2 (Franceschina, 2005; Sabourin et al., 2006).

En el marco de nuestro trabajo, nos inspiramos de estas observaciones procedentes del campo de la L1/2L1/L2 y de los modelos de L3 presentados en el capítulo 1 para guiar nuestro diseño experimental y formular nuestras hipótesis específicas sobre el proceso de adquisición del género gramatical en la L3. Como veremos más en detalle en la sección 2.6, el trío de lenguas de nuestro estudio nos permitirá explorar las posibilidades de influencia del fenómeno en el español L3, la cual, en nuestro caso, puede darse o bien a partir del francés, que presenta el un sistema de género gramatical equivalente al español, o bien a partir del euskera, que no materializa el género.

2.6 Diseño experimental

El hilo conductor de nuestras preguntas de investigación es el análisis de la influencia de las dos primeras lenguas sobre el español L3 en nuestros aprendientes bilingües (francés-euskera). En el caso del género gramatical, la única fuente de influencia que ofrece predicciones claras es el

francés, ya que comparte los mismos mecanismos de concordancia que el español y varios sustantivos equivalentes entre las dos lenguas comparten el mismo género. Con esto en mente, formulamos una primera pregunta preliminar (P1) que no concierne a la L3 sino una de las dos primeras lenguas. Para poder determinar de dónde procede la eventual influencia, hemos diseñado las tareas para permitir la recogida de datos en las tres lenguas de los participantes. Así, podremos comparar la L3 de un mismo participante con su propia L1 y L2 y no únicamente con una descripción teórica de sus lenguas. A partir de las respuestas obtenidas a esta pregunta preliminar, tendremos un punto de referencia fiable para analizar las mismas estructuras en la L3 de nuestros participantes y plantearnos una serie de preguntas adicionales (P2-4):

Pregunta 1: ¿Qué características muestran los SD producidos por nuestros participantes bilingües en francés en cuanto a la asignación y a la concordancia de género? ¿Son fieles a las descripciones que hemos podido observar en trabajos teóricos sobre el género gramatical en francés?

Pregunta 2: ¿Cuáles de los factores observados en adquisición de L2 provocan más errores de género en el SD en español L3?

Pregunta 3: ¿Cuáles de esos factores condicionan la influencia interlingüística (a partir de la L1 o L2) en español L3?

Pregunta 4: ¿Cómo evolucionan los efectos iniciales de la influencia interlingüística a lo largo del proceso de adquisición de español L3?

La pregunta 4 implica un seguimiento en el tiempo para observar la evolución dinámica de la L3 y es precisamente lo que motivó nuestro diseño longitudinal. Durante dos años, en tres momentos precisos, recogimos datos experimentales con los mismos participantes para observar cómo se desarrollaba su L3, en qué momento se ejercía la influencia y cómo variaba de un momento a otro (equivalentes a la progresión del nivel inicial al intermedio). La primera recogida de datos, después de una breve exposición a la lengua de menos de 30 horas de enseñanza (repartidas sobre aproximadamente dos meses), nos permitió observar las etapas iniciales del desarrollo de la lengua. Las recogidas sucesivas (después de 140 y 200 horas, respectivamente) nos permitieron observar qué cambios ocurrían en la interlengua de los participantes a medida que estos iban avanzando en su proceso de adquisición de la L3.

2.6.1 Hipótesis

En cuanto a la pregunta 1, las hipótesis son bastantes claras y corresponden a lo que hemos descrito en la sección 2.2.2 y, con ciertos matices, en las secciones 2.4.1 y 2.4.2. En francés, nos esperamos encontrar SD con una concordancia de género entre el sustantivo y sus modificadores, es decir, el determinante y el adjetivo. Si se observa una diferencia entre los dos tipos de concordancia, deberíamos observar más errores de concordancia con los adjetivos que con los determinantes, pero esa variación no debería ser importante dado que todos nuestros participantes tienen un nivel nativo o casi nativo en esta lengua.

El hecho de comprobar el nivel de dominio de la concordancia para todos los participantes nos permitirá considerar (o no) el francés como una posible fuente de influencia hacia la L3 en las hipótesis siguientes. En lo que respecta a la pregunta 2, consideramos que los participantes deberían reproducir de manera global lo observado en los estudios de L2 es decir: (2a) más aciertos con el género masculino (masculino por defecto), (2b) los determinantes (tipo de concordancia), (2c) las terminaciones transparentes (morfofonología) y (2d) el género equivalente al francés (criterio analógico).

En cuanto a la pregunta 3 y a la influencia interlingüística, según los modelos de transferencia de adquisición de L3 que hemos presentado en el capítulo 1, podemos postular una influencia o bien a partir del euskera, o bien a partir del francés. Según el *Typological Primacy Model* (TPM), el *Linguistic Proximity Model* (LPM) y el *Cumulative Enhancement Model* (CEM), la fuente de influencia debería ser el francés. En el trío de lenguas, el francés y el español son las dos lenguas tipológicamente más cercanas (TPM), los indicios (*micro-cues*) disponibles en cuanto a la concordancia de género en el SD deberían llevar a los aprendientes a establecer una similitud estructural entre el francés y el español (LPM) y la única influencia facilitadora o neutra puede proceder del francés (CEM). En cuanto al *L2 Status Factor*, la influencia debería proceder del francés solo si es la L2 de los participantes. En el caso contrario, se debería observar una influencia del euskera. Finalmente, en cuanto al llamado *L1 Factor*, la influencia debería proceder del francés solo si es la L1 de los participantes y del euskera en el caso contrario.

En concreto, la influencia a partir de las dos primeras lenguas merece más atención. A continuación, presentamos una serie de hipótesis sobre el género gramatical (G) sufijadas según si

la fuente de lengua hipotética es el euskera (EU) o el francés (FR).⁶⁷ Empezamos con el euskera, que, a pesar de que no materialice el género gramatical, presenta ciertas posibilidades de influencia. Nuestra primera hipótesis está relacionada con el efecto observado en estudios anteriores de L2 según si la L1 de los hablantes era [+gen] (francés, alemán, etc.) o [-gen] (inglés, euskera, etc.).

G1-EU: El euskera no materializa el género gramatical, por lo tanto, una influencia de esta lengua podría manifestarse con una tasa de concordancia más baja y, por lo tanto, una sobreutilización del masculino, el género por defecto.

Además, como menciona el estudio de Caffarra et al. (2017), los bilingües euskera-español dominantes de euskera parecen prestar más atención a la forma tanto en euskera como en español, por lo que es posible que la influencia del euskera en nuestros bilingües euskera-francés se concrete en una mayor atención a la forma en español, es decir, una mayor sensibilidad a la terminación de los sustantivos. Consideramos que esta mayor sensibilidad a la terminación de los sustantivos podría modular el efecto de la transparencia de manera positiva.

G2-EU: Si bien predecimos un efecto facilitador de la terminación transparente para todos los grupos, predecimos un efecto de mayor amplitud para los participantes que muestren una influencia interlingüística a partir del euskera.

En cuanto al francés, dado que los sistemas de género del español y del francés son similares, la influencia podría ser más directa.

G1-FR: Consideramos que los participantes que muestren una influencia del francés deberían beneficiarse de un efecto facilitador (mayor tasa de acierto de manera general), ya que los sistemas de género del español y del francés son similares.

G2-FR: La influencia del francés debería modular el efecto del criterio analógico, es decir, el género analógico debería facilitar la concordancia de manera más acentuada en los participantes que muestren una influencia del francés.

En cuanto a la pregunta 3, si bien esperamos que los resultados mejoren a lo largo del proceso de instrucción, como muestran los resultados de estudios de adquisición de L2 con aprendientes de nivel intermedio y avanzado, la evolución específica de cada factor de influencia interlingüística resulta más difícil de predecir, dada la escasez de estudios anteriores de L3 sobre nuestro objeto de estudio.

⁶⁷ Usaremos el mismo sistema de notación para las hipótesis del capítulo 3. Esto nos permitirá volver sobre las hipótesis de influencia interlingüística de manera más clara en el capítulo 4 sobre la discusión general.

2.6.2 Participantes

Como mencionamos en la introducción, nuestros participantes para los grupos experimentales procedían del País Vasco Norte y eran alumnos de dos centros educativos distintos: un colegio ubicado en la provincia de Baja Navarra ($n=26$) y otro ubicado en la provincia de Lapurdi ($n=28$). En estos centros educativos de la red Seaska, dado que la lengua de enseñanza es el euskera y que la lengua mayoritaria de la sociedad es el francés, los alumnos llegan a dominar sus dos primeras lenguas (el francés y el euskera) siendo muy jóvenes y, a los 10-11 años, pueden empezar los cursos de L3.

Al inicio del estudio, todos estaban en sexto grado del sistema de educación francés (11-12 años), habían escogido el español como L3 y no habían tenido contacto significativo con el español antes. Tenían cuatro horas de enseñanza de español en su horario semanal durante el primer año de las recogidas, y 3 horas semanales durante el segundo año. Todos los participantes tenían el euskera y el francés como lenguas previamente adquiridas, con variaciones en cuanto al orden y a la edad de adquisición de cada una de ellas. La frecuencia de uso y el nivel de dominio de cada lengua también variaba según los participantes.

Para agrupar a nuestros participantes, en lugar de solo basarnos en el orden de adquisición de las lenguas, preferimos considerar también el uso de la lengua como indicador de la lengua dominante del niño, siguiendo la clasificación del estudio de Caffarra et al. (2017) con bilingües español-euskera. Dado que, para todos, la lengua de instrucción era el euskera, con 4-5 horas de cursos de francés a la semana, consideramos el factor de la(s) lengua(s) hablada(s) en casa como el elemento diferenciador para dividirlos en tres grupos: los que hablaban francés y euskera (FR/EU), euskera (EU) o francés exclusivamente (FR). Para determinar la(s) lengua(s) hablada(s) en casa, nos basamos en las respuestas de los padres a las preguntas sobre 1) la lengua hablada con el padre y 2) la lengua hablada con la madre. Confirmamos estos datos con el niño mismo durante la entrevista que detallaremos a continuación. Dado que muchos de los niños hablaban las dos lenguas en casa el grupo FR/EU es el más numeroso. En los tres grupos experimentales, la edad de adquisición de la L2 o de la 2L1 variaba de un grupo a otro de manera significativa ($M=2.1$, $SD = 2.2$, $F(2,49) = 40.9$, $p < .001$), con un promedio particularmente alto en el grupo EU, lo cual indica

que se trata de participantes que habían adquirido su L2 en un contexto más formal (en el jardín de infancia o al inicio de la escuela primaria).⁶⁸

También contamos con un grupo de control de hablantes nativos de español. Este grupo procede de un colegio bilingüe (español-inglés) de Valladolid, en España. En el momento de la recogida, estos participantes estaban en su sexto curso de primaria (11-12 años). Elegimos un grupo de niños que habían recibido educación bilingüe para que sean más comparables con nuestros grupos del País Vasco, pero solo hemos recogido datos en español ya que se trata de nuestro grupo de control para el español exclusivamente.

La Tabla 2.3 presenta la información resumida sobre los participantes. Presentaremos en la próxima sección los instrumentos específicos que nos permitieron obtener estos datos.

⁶⁸ Para referirnos a las lenguas previas de nuestros participantes, en los casos de bilingüismo secuencial (grupos EU y FR), distinguimos entre L1 y L2 y, en el caso de bilingüismo simultáneo (grupo FR/EU), hablamos de 2L1.

Tabla 2.3. Resumen del perfil de los participantes según la(s) lengua(s) hablada(s) en casa.

	ES (n=21)	EU (n=14)	FR (n=15)	FR/EU (n=23)
Edad	11.4 (0.2)	11.3 (0.3)	11.4 (0.3)	11.3 (0.3)
Edad de adquisición L2 o 2L1	-	4.1 (1.8)	2.5 (1.8)	0.6 (1.3)
Horas de enseñanza en español	-	28 (0)	28 (0)	28 (0)
Prueba de nivel en español	117.6 (2.6)	12.8 (10.8)	12.5 (7.9)	9.8 (5.3)
Prueba de nivel en euskera	-	109.9 (8)	104.9 (6.8)	100.8 (11.1)
Prueba de nivel en francés	-	109.6 (10.5)	112.3 (4)	108 (7.9)

Nota. Los datos representan el valor promedio acompañado de la desviación estándar entre paréntesis. Los resultados de las pruebas de nivel son sobre 120.

Además de los datos biográficos de los participantes, en la Tabla 2.3 aparecen los resultados de una prueba de nivel que detallaremos a continuación. El objetivo de la prueba de nivel en consistía, por un lado, en confirmar el nivel inicial de los participantes en la L3 y, por otro, en confirmar la lengua dominante de los participantes (euskera o francés). Los tres grupos experimentales obtuvieron resultados bajos en la prueba en español, lo que nos permitió confirmar su nivel inicial, sin diferencia significativa de un grupo a otro ($M=11.5$, $SD = 8.15$, $F(2, 49) = 0.478$, $p = 0.63$). Estos resultados bajos son coherentes con la cantidad reducida de horas de enseñanza que habían recibido los participantes (28 horas). En cambio, el grupo de control obtuvo

un promedio cerca de la nota máxima ($M = 117.6$, $SD = 2.6$), lo que nos confirmó la validez de la prueba en español con niños de la misma edad.⁶⁹

2.6.3 Instrumentos

El primer instrumento que utilizamos para obtener información sobre nuestros participantes fue un cuestionario sobre su perfil sociolingüístico (completado por los padres de los participantes). A partir de esta información pudimos establecer cuáles eran los participantes que reunían las características que buscábamos.

Para completar el perfil de los participantes, utilizamos una prueba de nivel (de repetición elicitada) y una prueba de memoria fonológica a corto plazo (de repetición de dígitos).⁷⁰ Con respecto al fenómeno estudiado en este capítulo, una tarea escrita de denominación de imágenes nos permitió evaluar de manera experimental el desarrollo del género gramatical en las tres lenguas del estudio. Finalmente, una entrevista individual nos sirvió para confirmar los datos biográficos de los participantes. A continuación, presentamos cada instrumento en detalle.

2.6.3.1 Antecedentes lingüísticos y extralingüísticos

Antes de empezar el estudio, los padres de los participantes completaron un formulario que estaba disponible en dos formatos: electrónico, en la página web del estudio, o impreso (ver Anexo 1). Las preguntas del formulario tenían que ver con el perfil de los padres (lengua y lugar de origen de cada uno); información sobre las lenguas habladas por el participante, como su edad de adquisición, su nivel de dominio (estimado) y su exposición a las diferentes lenguas. Estos datos nos permitieron excluir a ciertos participantes que no cumplían con el perfil que buscábamos (por ejemplo, exposición anterior al español) y agrupar a todos los demás que sí correspondían al perfil buscado.

Confirmamos estos datos con cada participante durante una entrevista individual articulada entorno a una serie de preguntas sencillas e informales para establecer un contacto en español con

⁶⁹ Como veremos a continuación, también nos interesamos por la relación entre los resultados de la prueba de nivel y la de memoria fonológica a corto plazo realizada en la lengua dominante de cada participante y no observamos una correlación de Pearson significativa ($r = .39$ para los grupos experimentales, $r = .42$ para el grupo de control), basándonos en los mismos criterios de significatividad que Kim et al. (2016).

⁷⁰ La prueba de memoria a corto plazo nos sirvió también para comprobar la validez de la prueba de repetición elicitada.

el participante. Esta entrevista también nos permitió confirmar el nivel de lengua de los participantes, que evaluamos mediante la prueba que presentamos a continuación.

2.6.3.2 Pruebas diagnósticas

Considerando la edad de nuestros participantes, no era adecuado emplear una tarea como la adaptación del DELE desarrollada en McGill University que ha sido utilizada en numerosos estudios de adquisición de español L2 (Duffield et al., 2002; Duffield & White, 1999), ya que está destinada a adultos. Además, era necesario emplear una prueba relativamente breve, ya que queríamos evaluar el nivel de competencia de las tres lenguas de los niños. Con esto en mente, nos orientamos hacia una tarea de repetición elicitada ya que se trata de un tipo de tarea que cuenta con una larga historia en la medición de nivel de lengua tanto en L1 (Slobin & Welsh, 1967) como en L2 (Naiman, 1974), tanto con niños como adultos. Desde los principios de los años 2000, las pruebas de repetición elicitada han suscitado un cierto interés y numerosos estudios de L2 han revelado fuertes correlaciones entre los resultados de esas pruebas y los de pruebas estandarizadas (Tracy-Ventura et al., 2014). Más recientemente, este tipo de prueba ha sido utilizada específicamente como prueba diagnóstica de la lengua dominante en bilingües inglés-coreano (Lust et al., 2016). El principio general en el que se basa este tipo de prueba es que el hablante no puede repetir oraciones de una cierta extensión si no es capaz de analizarlas y entenderlas primero.

En el marco de nuestro estudio, utilizamos la prueba *Repeat as much as you can* (L. Ortega et al., 2002) para medir el nivel de dominio de las tres lenguas de los participantes. Se trata de una prueba en la que se presenta una serie de frases a un participante que tiene que escuchar e intentar repetirlas con la mayor exactitud posible. La prueba consta de 30 oraciones, de entre 7 y 18 sílabas cada una, que se presentan en orden de longitud y complejidad ascendente (de más corta a más larga y de más sencilla a más compleja). La duración de la prueba es de aproximadamente 10 minutos y requiere otros 10 minutos para ser puntuada. A la hora de evaluar, el calificador tiene que atribuir una nota de 0 a 4 a cada oración producida, logrando así una nota final que permite clasificar con precisión a los participantes. Cuando diseñamos el estudio, existían versiones de esta tarea en francés y en español, pero no en euskera.⁷¹ Por lo tanto, realizamos una adaptación

⁷¹La tarea existe también en coreano (Kim et al., 2016), alemán, japonés, chino-mandarín (L. Ortega et al., 2002) y ruso (Drackert, 2015).

completa para el euskera y, para las otras dos lenguas, tomamos las versiones existentes y aportamos una serie de modificaciones menores para adaptarlas a niños en lugar de adultos.

Siguiendo el diseño experimental de Kim et al. (2016), decidimos incluir también una tarea de memoria de dígitos para medir la memoria fonológica a corto plazo (MFCP). La combinación de esta tarea con la de repetición elicitada (*Repeat as much as you can*) nos permite asegurarnos de que esta última no mide la MFCP, sino el nivel de dominio de la lengua. Al evaluar la ausencia de correlación entre estos dos resultados, pudimos comprobar que la dificultad en repetir oraciones de la tarea anterior se debía a limitaciones de procesamiento de la lengua y no a limitaciones de la MFCP, una ausencia de correlación ya encontrada en Kim et al. (2016) con adultos y que encontramos también con los niños que participaron en nuestro estudio.

Para más información sobre el diseño y la lógica de la tarea, referimos al lector al trabajo reciente de Bowden (2016) o al trabajo original de Ortega et al. (2002) y, para todos los detalles sobre la adaptación de la tarea y el análisis detallado de los resultados obtenidos con nuestros participantes, referimos al lector a Lacroix y Alba de la Fuente (2019; en preparación).⁷²

2.6.3.3 Tarea de denominación de imágenes

Para estudiar el género gramatical, utilizamos una tarea de denominación de imágenes que incluía, por un lado, ítems para elicitación de la concordancia de género y, por otro, ítems para elicitación de compuestos. Nos concentraremos en los primeros en este capítulo y analizaremos los segundos en el capítulo siguiente. De los 68 ítems de la prueba, 36 estaban relacionados con el género gramatical, 24 tenían un color variable en cuanto al género (rojo, amarillo, blanco) y los 12 restantes comportaban un adjetivo de color invariable (verde, gris, azul). A los 68 ítems se suman 6 ítems de práctica (2 para cada tipo de compuesto y 2 para el género) para asegurarnos de que los

⁷² Si bien las pruebas en euskera y en francés nos permitieron comprobar que los niños tenían un alto nivel de dominio en sus dos primeras lenguas, ciertas limitaciones de la prueba nos impidieron basarnos exclusivamente en estos resultados para establecer la lengua dominante de los niños, por lo que, como comentamos antes, optamos por usar el criterio de la(s) lengua(s) usada(s) en casa para separar a los participantes por grupos. En resumen, las dos principales limitaciones eran las siguientes:

- Algunos ítems del vocabulario en francés no eran adaptados para los niños
- Ciertas estructuras en euskera parecen ser más difíciles que en otras lenguas para niños de la misma edad.

Referimos al lector a Lacroix y Alba de la Fuente (2019; en preparación) para más detalles sobre estas limitaciones.

participantes entendían bien la prueba y el tipo de respuesta esperada. La Figura 2.1 muestra dos ejemplos de ítems

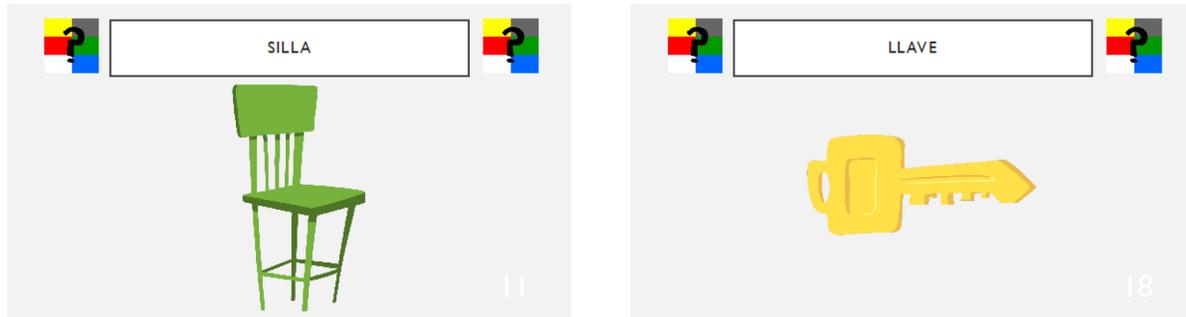


Figura 2.1. Ítems 11 y 18 de la prueba de denominación de imágenes en español.

En cada ítem, el participante debía observar la diapositiva y proporcionar una respuesta adecuada. Por ejemplo, para el ítem 11, que aparece en la Figura 2.1, la respuesta esperada era “una silla verde” y para el 18 “una llave amarilla”. En la prueba en español, como los participantes eran de nivel inicial, se les proporcionó el sustantivo junto con la imagen en la diapositiva (en estos casos “silla” o “llave”) para evitar respuestas en blanco. Además, dado que el objetivo de esta prueba consistía en evaluar el género gramatical y la concordancia a nivel del SD y no el simple conocimiento léxico, consideramos que el hecho de proporcionar el sustantivo no conllevaba problemas metodológicos.⁷³

Basándonos en los factores más frecuentes en los estudios anteriores sobre el género gramatical en L2, creamos ítems en función de las siguientes variables: género masculino (M) o femenino (F), terminación opaca (O) o transparente (T) y género analógico (A) o diferente (D) del equivalente en francés. Además de esos criterios, agregamos un criterio a nivel del ítem, que resulta de la combinación con un adjetivo variable (V) o invariable (I) en cuanto al género. Los ítems con un adjetivo invariable en cuanto al género eran considerados como distractores y se incluyeron para evitar que los participantes se acostumbraran a los adjetivos terminados en -o y -a. Para

⁷³ Además, el hecho de incluir el sustantivo en estos ítems cumplía un doble objetivo. Por un lado, minimizaba el riesgo de respuestas en blanco por desconocimiento del léxico y, por otro, nos permitió homogeneizar la estructura de los ítems en toda la prueba. Como veremos en el capítulo siguiente, los ítems sobre compuestos incluían texto e imágenes, de manera que, al incluir texto también en los ítems sobre género, sistematizamos la estructura de la prueba y logramos que estos ítems ejercieran con mayor eficacia la función de distractores de los ítems sobre compuestos y viceversa.

concentrarnos en las variables que dieron resultados más consistentes en los estudios anteriores, decidimos no incluir el género natural en nuestro diseño y, por lo tanto, todos los sustantivos escogidos representan una entidad inanimada. En la Tabla 2.4 aparecen los detalles de la distribución de los ítems según estos criterios. Para cada criterio, el número de ítems aparece entre paréntesis. En la última columna aparece un ejemplo de los cuatro criterios combinados.

Para establecer la lista de sustantivos, buscamos en estudios anteriores (Foote, 2014; Montrul et al., 2008) y en manuales de enseñanza del español como lengua extranjera (Borobio, 2011; Corpas, García, et al., 2013; López Alvarado, 2009; entre otros) para encontrar palabras que correspondieran a los criterios antes mencionados. Además, para evitar respuestas en blanco causadas por limitaciones en el vocabulario de los participantes, consideramos como criterio discriminatorio el grado de familiaridad, ilustrabilidad y concreción según el corpus ESPal (Duchon et al., 2013).⁷⁴ Solo los sustantivos con un promedio superior a 5/7 en estos tres criterios fueron considerados. El promedio global de la lista de sustantivos es de 5,68/7.

Tabla 2.4. Repartición de los ítems según las variables del estudio.

Género (ES)	Terminación	Género (FR)	Adjetivo	Ejemplo
M (18)	O (9)	A (6)	Variable (3)	Un corazón rojo
			Invariable (3)	Un árbol verde
		D (3)	Variable (3)	Un cohete rojo
	T (9)	A (6)	Variable (3)	Un libro rojo
			Invariable (3)	Un barco azul
		D (3)	Variable (3)	Un armario blanco
F (18)	O (9)	A (6)	Variable (3)	Una cruz blanca
			Invariable (3)	Una fuente azul
		D (3)	Variable (3)	Una nariz roja
	T (9)	A (6)	Variable (3)	Una casa roja
			Invariable (3)	Una botella gris
		D (3)	Variable (3)	Una cama roja

En la versión francesa de la tarea, los colores invariables en cuanto al género en español (“azul”, “gris” y “verde”) fueron reemplazados por invariables en cuanto al género en francés

⁷⁴ Las notas del corpus proceden de los juicios de hablantes nativos de español de distintas universidades españolas. Para cada criterio (familiaridad, ilustrabilidad y concreción), los participantes tenían que atribuir una nota en una escala del uno al siete. Cada palabra del corpus fue evaluada por un mínimo de 30 participantes.

(“*jaune*”, “*rouge*” y “*bleu*”).⁷⁵ En euskera, se conservaron los mismos ítems de la versión española, ya que, en esa versión no se aplica la concordancia de género. En las versiones en francés y en euskera, ya que los participantes tienen un nivel casi nativo o nativo, no aparece el sustantivo asociado a la imagen como en la versión en español para no sugerir una palabra concreta al participante. Por ejemplo, para el ítem que representa ‘un cohete’, los participantes, en francés, podrían pensar que se trata de “*une fusée*” o de “*un missile*”. De esta manera, evitamos una posible distorsión en los resultados, ya que el objetivo de las versiones en francés y euskera de la prueba era el de obtener la palabra asociada a la imagen en la mente del hablante y no una hipotética traducción del diccionario.

2.6.4 Protocolo de administración de las tareas

Como mencionamos antes, el primer paso en la recogida de datos era el cuestionario que tenían que completar los padres. Después, se administraron todas las tareas en español, en el siguiente orden: tarea de denominación de imágenes, prueba de nivel y entrevista individual. Estas tareas se administraron en dos sesiones de trabajo separadas dentro de una misma semana.

La tarea de descripción de imágenes fue administrada en una clase equipada con un proyector en la cual todos los participantes de una misma clase (entre 12 y 20 participantes) realizaron la tarea simultáneamente y de manera individual. Antes de empezar, el investigador explicó la tarea a los participantes y les entregó una hoja de respuesta individual y un mini-diccionario visual (ver Anexo 2). Las explicaciones se dieron en el idioma de la tarea, pero, dado que los participantes tenían un nivel bajo de dominio del español, el investigador y el docente contestaron a las preguntas en francés o en euskera cuando resultó necesario para aclarar las instrucciones de la tarea. Además de las explicaciones, el investigador repasó los seis colores que los participantes tenían que usar para responder (“rojo/a”, “blanco/a”, “amarillo/a”, “azul”, “gris” y “verde”). Después, el investigador prosiguió con los seis ítems de práctica y escribió las respuestas en la pizarra. Antes de empezar, el investigador explicó que cada ítem aparecería durante solo 25 segundos en la pantalla y que, por lo tanto, tenían que escribir la primera respuesta que les viniera a la mente. Se insistió también en que las respuestas eran individuales y en que era importante escribir una respuesta para cada diapositiva. A la mitad de la tarea, se realizó una pausa

⁷⁵ Si bien el adjetivo “*bleu*” varía en género, su femenino solo se diferencia por una –e final (muda).

de 2 minutos. La tarea desde el inicio de las explicaciones hasta el final duró entre 35 y 45 minutos, según la duración de las explicaciones.

En otra sesión de trabajo, los participantes realizaron de manera individual la prueba de nivel en un ordenador equipado con audífonos y micrófono, bajo la supervisión del investigador. La entrevista individual se realizó inmediatamente después de la prueba de nivel, en la misma sesión de trabajo.

Una semana más tarde (como mínimo), se administraron dos de las pruebas en euskera siguiendo el mismo orden: denominación de imágenes y prueba de nivel, en dos sesiones de trabajo distintas. Una semana más tarde, se repitió la misma operación con las pruebas en francés. Por último, se realizó una entrevista individual en las dos primeras lenguas del participante (en la L2 y en la L1, una tras otra) y la prueba de memoria de dígitos en la lengua dominante del participante.

Las tareas en euskera y en francés solo se administraron en la primera etapa de recogida de datos (después de 28 horas de enseñanza en español). En cuanto al español, para ver cómo evolucionaba el dominio del fenómeno en español, repetimos el mismo procedimiento dos veces más. Usando las mismas tareas, pudimos observar la evolución de cada participante después de 140 horas de enseñanza en español y una vez más después de alrededor de 200 horas.

2.7 Resultados

Antes de comprobar nuestras hipótesis con los datos de español como L3, empezamos con la presentación de nuestros datos de francés y de euskera para contestar a nuestra primera pregunta de investigación, es decir, para asegurarnos de que las respuestas de nuestros participantes en francés se corresponden con las respuestas esperadas (en cuanto al género asignado). Además, para comprobar la adecuación de la tarea que nos sirvió a elicitar los diferentes SD, presentaremos los resultados de nuestro grupo de control de español antes de presentar los resultados de español L3.

En nuestros análisis estadísticos, incluimos todos los factores de los ítems (género, terminación, analogía con el francés; ver Tabla 2.4) y los factores individuales (grupo). Para todos los ítems, la concordancia del sustantivo con el determinante y con el adjetivo fueron codificadas como dos respuestas separadas, por lo que también incluimos el tipo de concordancia en nuestros análisis. Codificamos la concordancia de manera binaria (0=error, 1=concordancia), lo que nos permitió analizar las respuestas con varios modelos de regresión logística de efectos mixtos

mediante el paquete *lme4* (D. Bates et al., 2015), en el programa *R* (R Core Team, 2018). En todos los casos, incluimos las estructuras aleatorias máximas, salvo cuando el nivel de complejidad del modelo impedía su convergencia. Cuando los modelos no convergían, eliminamos, en primer lugar, la correlación entre las pendientes (*slopes*) y los interceptos (*intercept*) aleatorios, las interacciones en las pendientes aleatorias, luego las pendientes aleatorias y, por último, las interacciones de mayor nivel entre los efectos fijos, conservando, como mínimo, los interceptos de sujeto y de ítem. La inclusión de estos efectos aleatorios, como recomiendan Baayen et al. (2008), permite tomar en cuenta la variabilidad aleatoria en nuestros datos experimentales. Para cada respuesta, tras el ajuste de un primer modelo que convergía, simplificamos el modelo eliminando los efectos uno por uno, guiándonos por el resultado de la función *dredge* de *MuMIn* (Bartoń, 2019), para seleccionar el modelo que tenía la mejor bondad de ajuste, según su AIC (*Akaike Information Criterion*). Por lo tanto, para cada respuesta, el modelo resultante solo incluía los efectos que mejoraban las predicciones globales del modelo, sin aumentar innecesariamente su complejidad.

2.7.1 Francés L1/2L1/L2

En francés, tenemos los datos de los 52 participantes para 36 ítems, lo que nos da un total de 1872 respuestas que contienen a su vez dos elementos de concordancia (determinante y adjetivo). En la presente sección, a diferencia de nuestro análisis de la prueba en español L3, contemplamos los 36 ítems, pero solo analizamos el factor de género, ya que el criterio analógico (francés/español) y la transparencia de la terminación (en español) no deberían afectar las respuestas en francés. Incluimos también en nuestro análisis el tipo de concordancia (con el determinante o el adjetivo). Antes de entrar en el análisis de la concordancia de género, si bien no analizamos el criterio analógico como posible efecto sobre la concordancia de género en francés, presentamos un análisis del género de las respuestas proporcionadas que nos servirá de base para analizar el criterio analógico en los datos de español L3.

Como lo comentamos en la sección 2.6.3.3, en la tarea en español L3, proporcionamos el sustantivo a los participantes para evitar respuestas en blanco, pero no era así en la tarea en francés (y tampoco en euskera o en español L1) porque queríamos averiguar a qué sustantivo correspondía la imagen en la lengua nativa o casi nativa de los participantes. Recordamos al lector que este elemento es de gran importancia para analizar el impacto del criterio analógico en español L3. De todas las respuestas proporcionadas, apenas unas 55 correspondían a un sustantivo alternativo,

como en (2.21) y, de estas, solo 28 correspondían al género opuesto al género esperado en francés, como en (2.22). En los ejemplos a continuación, como en todas las respuestas alternativas que presentaremos a lo largo de la sección de resultados, el sintagma después de la barra diagonal representa la respuesta esperada.

- (2.21) a. *un cartable blanc / un sac à dos blanc*
 un mochila.MASC blanco / un mochila.MASC blanca
- b. *une jupe blanche / une robe blanche*
 una falda blanca / una vestido.FEM blanca
- (2.22) a. *une chemise blanche / un t-shirt blanc*
 una camisa blanca / un camiseta.MASC blanco
- b. *un placard blanc / une armoire blanche*
 un armario blanco / una armario.FEM blanca

Ya que estas últimas respuestas representan el 1,5 % de los datos y están repartidas sobre 7 ítems diferentes, para simplificar nuestro análisis en español L3, partiremos del presupuesto que el género analógico del francés para los sustantivos en español es el que proponemos en nuestro diseño experimental y no el género expresado en las pocas respuestas divergentes.

En todo el juego de datos, tenemos solo 6 respuestas en blanco o ilegibles, pero encontramos un número considerable de omisión de determinante ($n=116$) en las respuestas y unos pocos casos de determinantes invariables en cuanto al género ($n=2$), es decir, los indefinidos “*des*” y “*du*”.⁷⁶ Con respecto a los adjetivos, encontramos 45 omisiones de adjetivos o respuestas que contienen un modificador alternativo, como mostramos en los ejemplos (2.23). En cambio, el número de adjetivos invariables en cuanto al género es más considerable ($n=620$), ya que representaba la respuesta esperada para 12 ítems por participante, es decir, 624 respuestas potenciales. La diferencia entre los adjetivos invariables esperados y los adjetivos invariables producidos se debe al hecho de que ciertos participantes cambiaron el nombre del color de los objetos, como mostramos en los ejemplos (2.24).⁷⁷ Además, encontramos 5 respuestas para las cuales no hemos

⁷⁶ Las omisiones de determinantes están repartidas entre 18 participantes, pero dos participantes contribuyen de manera más clara, ya que omitieron los determinantes en todas sus respuestas.

⁷⁷ Los participantes confundieron a menudo los colores gris y azul porque, con el proyector que nos sirvió para presentar los estímulos, el contraste de color no era muy claro. Cabe aclarar aquí que, si bien “*bleu*” (“azul”) es variable en francés (“*bleu*” en masculino, “*bleue*” en femenino), esta diferencia no es perceptible en la lengua oral y podría

podido identificar el género del sustantivo, por ser un sustantivo tomado prestado del euskera, como en (2.25) o un sustantivo incomprensible, como en los ejemplos (2.26).

- (2.23) a. *une fontaine à eau* / *une fontaine bleue*
“una fuente de agua” / “una fuente azul”
b. *une moustache de vieux* / *une moustache jaune*
“un bigote de viejo” / “un bigote amarillo”
- (2.24) a. *un jean gris* / *un pantalon bleu*
un pantalón gris / un pantalón azul
b. *un nuage bleu* / *un nuage gris*
una nube azul / una nube gris
- (2.25) a. *la iturria bleu* / *une fontaine bleue*
la fuente azul / una fuente azul
b. *un dorre* / *une tour grise*
un torre / una torre gris
- (2.26) a. *une sobine bleu* / *une fontaine bleue*
una ????? azul / una fuente azul
b. *un lure blanc* / *un mur blanc*
un ??? blanco / una pared blanca

Una vez excluidos estos datos en los cuales no era posible analizar la concordancia con el sustantivo, empezamos nuestro análisis con un juego de datos de 2939 entradas (1745 determinantes, 1194 adjetivos).

De manera global, la tasa media de concordancia es del 93,16 %, con una tasa más alta con los determinantes (95,13 %) que con los adjetivos (90,28 %). En cuanto a la diferencia en función del género del sustantivo, la tasa de concordancia asciende al 97,68 % con el masculino y baja al 88,66 % con el femenino. Con respecto a los grupos de participantes, el grupo FR tiene la tasa más alta con un 95,19 % de concordancia, seguido por los grupos FR/EU (93,36 %) y EU (90,76 %). La Figura 2.2 ilustra las tasas de concordancia media para cada grupo en cada condición.

representar un desafío ortográfico incluso para los monolingües de la edad de nuestros participantes y no ser un fiel reflejo de la concordancia de género. Entonces, hemos considerado este color como invariable para nuestros análisis.

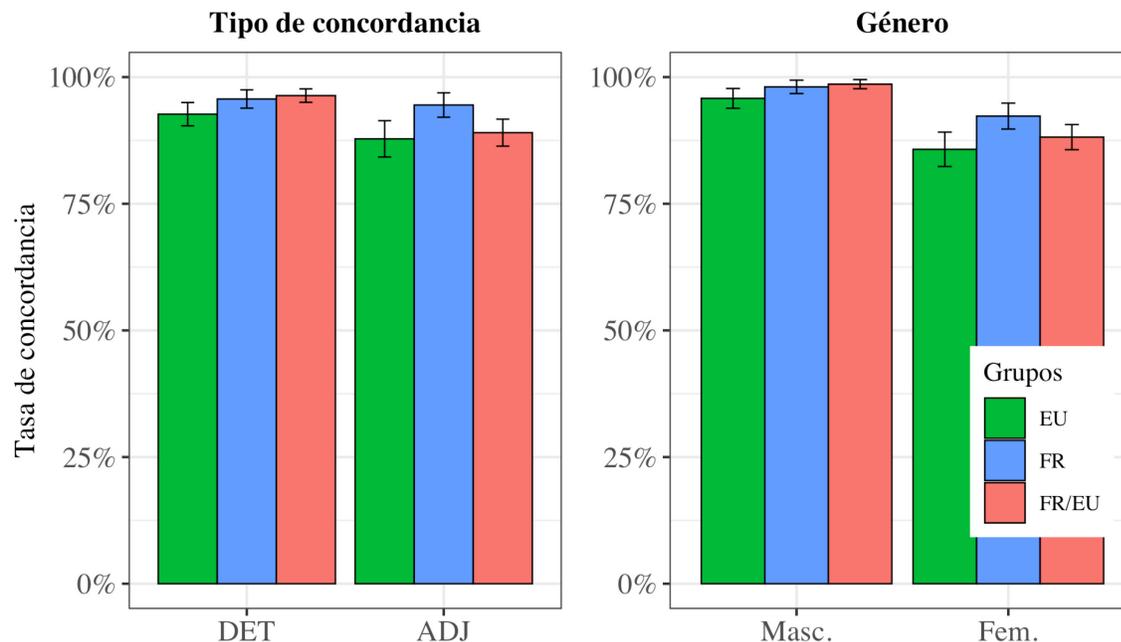


Figura 2.2. Tasa de concordancia de género por grupo y condición en la prueba de control en francés.

Si bien las tasas de concordancia son relativamente altas en todas las condiciones, se observan variaciones. Para evaluar los efectos de grupo, tipo de concordancia y género sobre los aciertos (1=concordancia; 0=error), construimos un modelo de regresión logística de efectos mixtos, en el que incluimos estas variables como efectos fijos y también la posible interacción entre ellas: grupo \times tipo, grupo \times género y tipo \times género. En cuanto a la estructura de los efectos aleatorios, incluimos un intercepto aleatorio por participante e ítem, y una pendiente aleatoria por participante para el tipo de concordancia. Los factores fueron codificados de la siguiente manera (el nivel de referencia corresponde al primer nivel de codificación): grupo (1=FR, 2=FR/EU, 3=EU), tipo (-.5=determinante, .5=adjetivo), y género (-.5=masculino, .5=femenino).⁷⁸ El proceso

⁷⁸ La centralización del contraste de los factores categoriales de dos niveles (-.5 vs .5) se utiliza para facilitar la interpretación del modelo y evitar la colinealidad de los múltiples efectos en los modelos de regresión. En cuanto al factor grupo, ordenamos los niveles siempre en función de lo que era más relevante para la interpretación de los modelos. En este primer modelo, dado que pensábamos observar una mayor diferencia entre el grupo FR y los dos otros grupos, utilizamos el grupo FR como nivel de referencia. En los siguientes modelos, siempre utilizaremos el grupo de control (ES) como nivel de referencia cuando sea aplicable, o el grupo EU cuando apliquemos un modelo únicamente a los datos de los grupos experimentales. La elección del grupo EU como nivel referencia se debe a que

de eliminación regresiva por paso (*stepwise backward elimination*) nos reveló que todos los efectos y sus interacciones contribuían a la bondad de ajustes del modelo, por lo que presentamos en la Tabla 2.5 la lista completa de los efectos fijos y de sus coeficientes respectivos.

Tabla 2.5. Resultados del modelo regresión logística de efectos mixtos de la concordancia de género en francés

	β	<i>SE</i>	<i>z</i>	<i>Pr(> z)</i>
(Intercept)	4.194	0.487	8.609	<.001
Grupo (FR/EU)	-0.020	0.562	-0.036	.971
Grupo (EU)	-0.743	0.598	-1.242	.214
Género (FEM)	-2.119	0.579	-3.660	<.001
Tipo (ADJ)	0.343	0.516	0.664	.507
Grupo (FR/EU) × Género (FEM)	-0.961	0.589	-1.631	.103
Grupo (EU) × Género (FEM)	0.150	0.530	0.283	.777
Grupo (FR/EU) × Tipo (ADJ)	-1.285	0.563	-2.282	.022
Grupo (EU) × Tipo (ADJ)	-0.423	0.590	-0.717	.474
Género (FEM) × Tipo (ADJ)	-1.044	0.459	-2.275	.023

Este modelo nos revela que, si bien los grupos FR y FR/EU producen más errores que el nivel de referencia (grupo FR), las diferencias entre los grupos no son significativas por sí solas. Por su parte, el efecto de género (femenino) es significativo y afecta de manera negativa a todos los grupos. El efecto tipo (adjetivo) no es significativo por sí solo, pero cuando interactúa con otras variables sí lo es. Así, el efecto de la interacción grupo (FR/EU) × tipo (ADJ) es significativo, por lo que, si bien la diferencia entre los grupos no es significativa de manera global, el grupo FR/EU produce significativamente más errores de concordancia con los adjetivos que con los determinantes, con una diferencia mayor que los dos otros grupos. Por último, el efecto Género (FEM) × Tipo (ADJ) también es significativo, también con un coeficiente negativo, es decir, que los participantes producen significativamente más errores de concordancia cuando el adjetivo tiene que concordar en femenino que con el determinante en masculino.

Ahora, al mirar de manera más detallada la condición en la que se producen más errores (adjetivo, femenino, $n=99$), constatamos que la gran mayoría de los errores ocurren con dos

este representaba el mayor contraste frente a los demás grupos, lo cual facilitaba la interpretación de los resultados. En cualquier caso, una batería de análisis preliminares nos permitió observar que el cambio de nivel de referencia no afectaba la significatividad de los efectos observados.

adjetivos: “gris” (50,50 %) y “vert” (26,26 %). En cambio, el adjetivo “blanc” solo aparece en 12,22 % de los errores en femenino.⁷⁹ Contemplar estos errores a la luz de las diferencias fonológicas entre el francés y el euskera es de particular interés en este caso. El francés, como comentamos en la sección 2.2.2 es una lengua que tiene una ortografía opaca, es decir, cada letra o combinación de letras puede corresponder a varios fonemas distintos y cada fonema puede corresponder a varias letras o combinaciones de letras diferentes. En el caso que nos interesa aquí, las consonantes que aparecen a finales de palabras pueden ser mudas en francés, y solo se pronuncian cuando van seguidas de una vocal, como es el caso cuando se forma el femenino de los adjetivos que nos interesan aquí. Así, en masculino, “gris” y “vert” se pronuncian respectivamente /gri/ y /vɛr/ y sus formas femeninas “grise” y “verte” se pronuncian /griz/ y /vɛrt/. En cambio, el euskera tiene una ortografía más transparente, como el español, por lo que “gris” se pronuncia /gris/ y la hipotética palabra “vert” se pronunciaría /vert/. Considerando esto, es posible que ciertos de esos errores estén relacionados con la ortografía opaca del francés más que a errores de concordancia como tal. Sin embargo, como se trata de un detalle que no habíamos planeado en nuestro diseño, los pocos errores no nos permiten hacer un análisis estadístico más detallado del fenómeno.

Volviendo a las respuestas de manera más global, considerando las tasas relativamente altas de respuestas esperadas, podemos considerar que la prueba de control en francés cumplió con su objetivo, que consistía en elicitación de sintagmas determinantes formados por un determinante, un sustantivo y un adjetivo de color. Además, si bien se observó una pequeña diferencia entre los grupos en la concordancia con los adjetivos, con una tasa de aciertos algo inferior por parte del grupo FR/EU en comparación con los otros dos grupos, no observamos diferencias entre los grupos de manera global; todos parecen dominar los mecanismos de concordancia de género en francés.

2.7.2 Euskera L1/2L1/L2

Como mencionamos en la sección 2.2.3, en euskera, no se materializa el género gramatical. A pesar de ello, como los participantes también completaron la tarea en euskera, reportamos de

⁷⁹ Los demás errores corresponden a formas idiosincrásicas como “blan” (4,08 %), “gri” (2,04 %), “ver” (5,10 %), formas que no constan de la consonante final. Los tres colores (gris, verde y blanco) aparecían cada uno con un tercio de los ítems, por lo que hubiera sido más probable obtener un tercio de los errores con cada adjetivo.

manera breve los datos recopilados para verificar que las respuestas corresponden a la forma esperada.

Tenemos los datos de los 52 participantes para 36 ítems, lo que nos da un total de 1876 respuestas. De estas respuestas, el 98,07 % corresponde a la estructura esperada, es decir, un sustantivo seguido por un adjetivo de color y un determinante ($n=1795$), como los ejemplos en (2.27), o simplemente un sustantivo seguido por un adjetivo, sin determinante, como en (2.28) ($n=45$). Las 36 respuestas inesperadas consisten en un sustantivo sin adjetivo ($n=20$), como en (2.29), o respuestas en blanco ($n=16$).⁸⁰

- (2.27) a. *bihotz gorria*
corazón rojo.det
“el corazón rojo”
b. *giltza hori bat*
llave amarilla una
“una llave amarilla”
- (2.28) a. *mahai berde*
mesa verde
“mesa verde”
b. *intzaur gris*
nuez gris
“nuez gris”
- (2.29) a. *luzokerra*
pepino.det
“el pepino”
b. *#coboy chapeaua*
vaquero sombrero-el
“sombrero de vaquero”

⁸⁰ A pesar de contener dos préstamos con una grafía impropia del euskera, el ejemplo (2.29b) se puede entender como un compuesto N+N bien formado, con el núcleo a la derecha. Es posible que respuestas como esta hayan surgido como un efecto de la tarea, ya que esta incluía ítems para elicitación 1) sintagmas determinantes formados por un sustantivo modificado por un adjetivo, 2) compuestos N+N y 3) compuestos deverbales. Los dos tipos de compuestos se estudiarán en el próximo capítulo. También encontramos en los datos respuestas híbridas que consisten en un compuesto verbal seguido por un adjetivo de color, como en (a), que hemos codificado como un sustantivo seguido por un adjetivo y un determinante.

- a) *#ur- botatzaile urdina / iturri urdina*
agua- botador azul.det / fuente azul.det
“el botador de aguas azul” / “la fuente azul”

Como las respuestas inesperadas representan apenas el 1,92 % de las respuestas, consideramos que la tarea cumplió su papel también en euskera.

2.7.3 Español L1

En el grupo de control, tenemos los datos de los 21 participantes para 36 ítems, lo que nos da un total de 756 respuestas que contienen a su vez dos elementos de concordancia (determinante y adjetivo). Al contrario de nuestro análisis de la prueba en español como L3, contemplaremos todos los 36 ítems, pero solo analizamos el factor de género y la transparencia de la terminación (en español), ya que el criterio analógico (francés/español) no debería afectar las respuestas porque estos participantes no hablan francés. Incluimos también en nuestro análisis el tipo de concordancia (con el determinante o el adjetivo).

En todo el juego de datos, tenemos solo 4 respuestas en blanco o ilegibles, pero encontramos algunas omisiones de determinante ($n=22$) y pocas omisiones de adjetivos ($n=7$) en las respuestas. En cambio, el número de adjetivos invariables en cuanto al género es más considerable ($n=195$), ya que representaba la respuesta esperada para 12 ítems por participante, es decir, 252 respuestas potenciales. La diferencia entre los adjetivos invariables esperados y los adjetivos invariables producidos se debe al hecho de que ciertos participantes cambiaron el nombre del color de los objetos, como ilustramos en los ejemplos (2.30).

- (2.30) a. una torre negra / una torre gris
b. un bigote verde / un bigote amarillo

Una vez excluidos estos datos, en los cuales no era posible analizar la concordancia con el sustantivo, empezamos nuestro análisis con un juego de datos de 1280 entradas (730 determinantes, 550 adjetivos).

De manera global, la tasa media de concordancia es del 98,59 %, con una tasa ligeramente más alta con los determinantes (98,76 %) que con los adjetivos (98,36 %). En cuanto a la diferencia en función del género del sustantivo, la tasa de concordancia asciende al 99,36 % con el masculino y baja al 97,87 % con el femenino. En cuanto a la transparencia, la tasa de concordancia es de 99,36 % en la condición transparente y del 97,85 % en la condición opaca. Con tasas de concordancia tan altas, resulta imposible analizar la variación, ya que está casi ausente. Entonces,

podemos concluir que la tarea en español también cumplió con su objetivo, que consistía en elicitación de sintagmas determinantes formados por un determinante, un sustantivo y un adjetivo de color. No se observó ningún patrón de respuestas inesperadas en particular, por lo que pasamos a los resultados de los grupos experimentales.

2.7.4 Español L3: las etapas iniciales

Como vimos en las subsecciones anteriores, la tarea elicitó principalmente la estructura esperada en las lenguas nativas o casi nativas de todos los participantes. Concretamente, en francés y en español, las dos lenguas que materializan el género gramatical, los participantes obtuvieron tasas elevadas de concordancia. A modo recapitulativo, la Tabla 2.6 resume las tasas de concordancia obtenidas en la tarea según el género esperado y el tipo de concordancia en las dos lenguas.

Tabla 2.6. Tasa de concordancia en francés (L1, L2 o 2L1) y en español (L1)

	Francés	Español
Género masculino	97,68 %	99,36 %
Género femenino	88,66 %	97,87 %
Concordancia con el determinante	95,13 %	98,76 %
Concordancia con el adjetivo	90,28 %	98,36 %
Promedio de concordancia global	93,16 %	98,59 %

Si bien todas las tasas presentadas en la Tabla 2.6 son elevadas, se puede observar que, en español, los efectos de género y de tipo de concordancia son menos perceptibles que en francés. Nuestro diseño experimental no nos permite comprobar si esta diferencia entre las dos lenguas se debe a una dificultad intrínseca de la lengua (el francés es más opaco) o si se debe al perfil de nuestros grupos experimentales.⁸¹ A pesar de ello, estos porcentajes son lo suficientemente altos

⁸¹ Concretamente, para desambiguar entre estas dos posibilidades, se necesitaría un grupo de control de niños monolingües de francés de la misma edad.

como para demostrar que los grupos experimentales no presentan problemas con la concordancia de género en francés.

Ahora, en cuanto a los datos en español como L3, descartamos los 12 ítems con adjetivos invariables para no desequilibrar nuestro juego de datos (para no tener más ítems con el género analógico del francés). Por lo tanto, tenemos los datos de los 52 participantes experimentales, para 24 ítems, lo que nos da un total de 1248 respuestas que contienen a su vez dos elementos de concordancia (determinante y adjetivo). A diferencia de los datos en las lenguas nativas o casi nativas, encontramos numerosas omisiones de determinantes y adjetivos en las respuestas de los participantes, por lo que tuvimos que analizar los patrones de omisión para limpiar los datos antes de empezar el análisis. Encontramos 382 respuestas sin determinantes y 131 respuestas sin adjetivos, repartidas de manera heterogénea entre los participantes. Para identificar los patrones de omisión de los participantes, calculamos el promedio de omisión y la desviación estándar por participante y por tipo de concordancia e identificamos los participantes que se alejaban de más de 1,5 desviaciones estándar del promedio de omisión. Así, encontramos 11 participantes que correspondían a este criterio para los determinantes y 5 para los adjetivos.⁸² Dado que no encontramos patrones de omisión similares en nuestras tareas en las lenguas nativas o casi nativas y tampoco en nuestras recogidas posteriores, atribuimos estas omisiones a una posible incompreensión de las instrucciones de la tarea. Entonces, para no distorsionar nuestro análisis con estos datos de omisiones, decidimos quitar todas las respuestas de estos participantes por tipo de concordancia, es decir, 264 respuestas para los determinantes y 120 para los determinantes. Hecho esto, excluimos las 83 omisiones restantes (35 determinantes, 48 adjetivos) ya que tampoco era posible analizar la concordancia con el sustantivo en estas respuestas. A continuación, desarrollamos nuestro análisis con un juego de datos de 2029 entradas (949 determinantes, 1080 adjetivos).⁸³

De manera global, la tasa media de concordancia es del 55,88 %, con una tasa más alta con los determinantes (59,43 %) que con los adjetivos (52,77 %). En cuanto a la diferencia en función

⁸² Ninguno de los participantes omitió a la vez todos los determinantes y todos los adjetivos. Por los 52 participantes entran en nuestro análisis.

⁸³ Con esta limpieza selectiva de los datos, conservamos un máximo de datos que contenían una respuesta analizable (1=concordancia; 0=error), sea en un determinante o en un adjetivo.

del género del sustantivo, la tasa media de concordancia asciende al 81,62 % con el masculino y baja al 30,28 % con el femenino. Cuando combinamos los dos factores, observamos que, en la concordancia entre un sustantivo de género femenino y un adjetivo, la tasa de acierto es de apenas 16,08 %. A nivel individual, 16 participantes muestran una tasa de concordancia de 0 % con los adjetivos en la condición femenina.⁸⁴ Con respecto a la transparencia, la tasa media de concordancia asciende al 59,72 % con los sustantivos de terminación transparente y baja al 52,06 % con los de terminación opaca. En función del criterio analógico, la tasa media de concordancia es del 63,97 % con los sustantivos cuyo género es congruente con su equivalente en francés y del 47,60 % en la condición incongruente.

Con respecto a los grupos de participantes, el grupo EU tiene la tasa más alta de concordancia con un 62,23 %, seguido por los grupos FR (56,07 %) y FR/EU (51,91 %). Estos resultados contrastan con los obtenidos con el grupo de control en español L1 (99,07 % de concordancia con los 24 ítems analizados en la presente sección). Antes de continuar con el análisis de los factores, para confirmar la significatividad de esta diferencia entre el grupo de control y los grupos experimentales, construimos una regresión logística de efectos mixtos en la cual incluimos el efecto fijo de grupo (1=ES, 2=EU, 3=FR, 4=FR/EU) y una pendiente aleatoria por participante y por ítem.⁸⁵ La diferencia es significativa en el contraste entre todos los grupos experimentales con el grupo de control: EU – ES ($\beta = -4.877$, $SE = 0.369$, $z = -13.20$, $p < .001$), FR – ES ($\beta = -5.239$, $SE = 0.371$, $z = -14.11$, $p < .001$) y FR/EU – ES ($\beta = -5.485$, $SE = 0.366$, $z = -15.00$, $p < .001$). En cambio, entre los grupos experimentales, solo el contraste entre FR/EU – EU es significativo ($\beta = -0.608$, $SE = 0.141$, $z = -4.324$, $p < .001$), mientras que el contraste FR – EU es marginalmente significativo ($\beta = -0.363$, $SE = 0.156$, $z = -2.318$, $p = .084$) y el contraste FR/EU – FR no lo es ($\beta = -0.246$, $SE = 0.138$, $z = -1.781$, $p = .261$).

Una vez confirmada la significatividad de las diferencias entre el grupo de control y los grupos experimentales a nivel global, nos concentramos exclusivamente en los datos de los grupos

⁸⁴ Si bien varios de estos participantes pertenecen al grupo FR/EU, como veremos más adelante, la interacción grupo \times tipo \times género no es significativa, por lo que consideramos que se trata de una variación individual.

⁸⁵ La simplicidad de este modelo se justifica por el hecho de que, por una parte, no podemos incluir factores que técnicamente no pueden afectar al grupo de control (como la congruencia) y, por otra, no podemos incluir factores que muestran muy poca variación en los resultados del grupo de control (transparencia, género y tipo de concordancia). Todos estos factores sí están incluidos en nuestro análisis de los grupos experimentales.

experimentales. La Figura 2.3 ilustra las tasas medias de concordancia para cada grupo experimental en cada una de las condiciones que analizaremos a continuación.

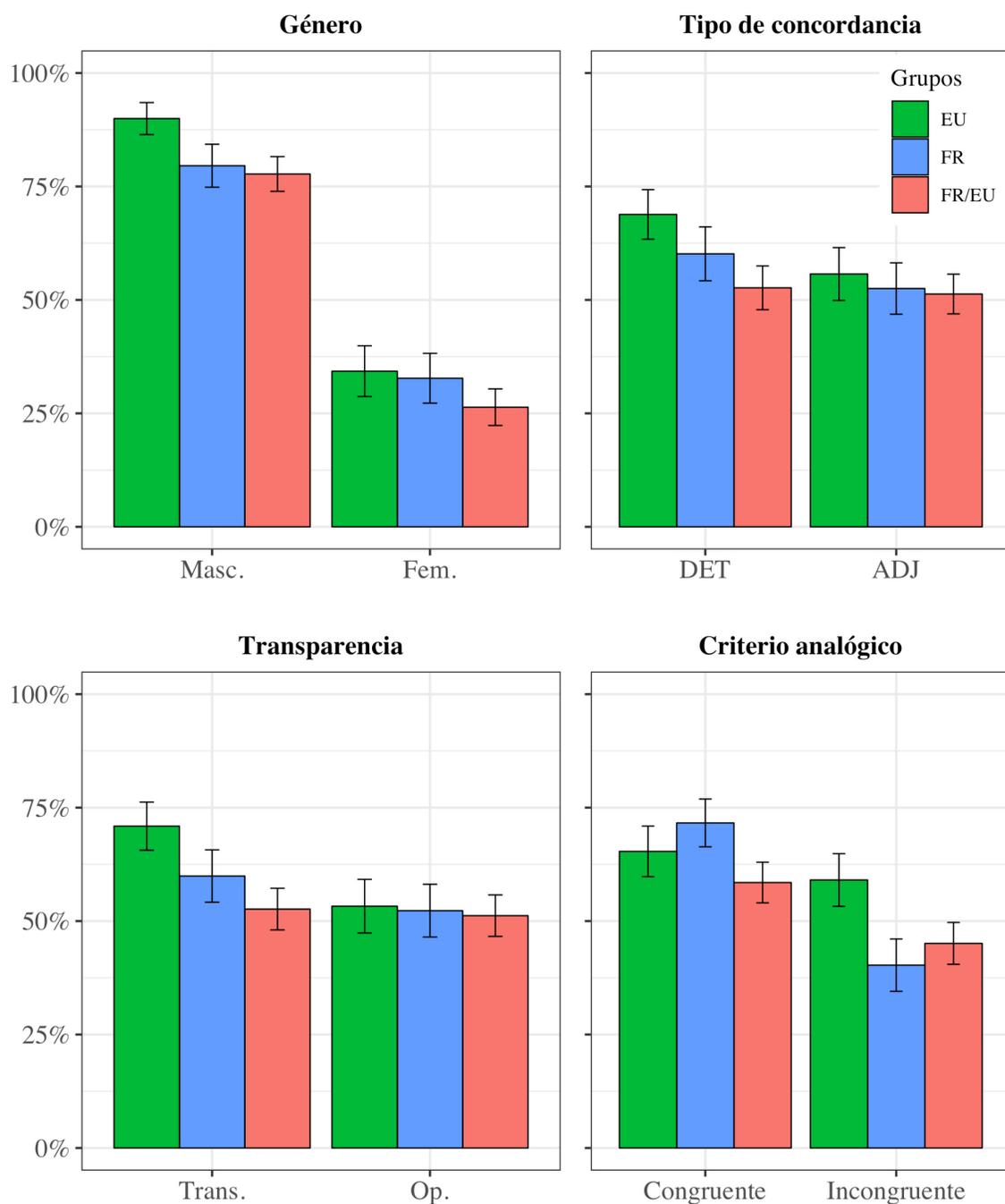


Figura 2.3. Tasa media de concordancia por grupo según las diferentes condiciones analizadas.

Para evaluar los efectos de grupo, género, tipo de concordancia, transparencia y congruencia sobre los aciertos (1=concordancia; 0=error), elaboramos un modelo de regresión logística de efectos mixtos, en el que incluimos estas variables como efectos fijos y también la posible interacción entre ellas grupo × congruencia, grupo × transparencia y tipo × género.⁸⁶ En cuanto a la estructura de los efectos aleatorios, incluimos un intercepto aleatorio por participante e ítem, y una pendiente aleatoria por participante para el género, la transparencia y la congruencia. Los factores fueron codificados de la siguiente manera (el nivel de referencia corresponde al primer nivel de codificación): grupo (1=EU, 2=FR, 3=FR/EU), género (-.5=masculino, .5= femenino), tipo (-.5=determinante, .5=adjetivo), transparencia (-.5 = opaco, .5=transparente) y congruencia (-.5=incongruente, .5=congruente). El proceso de eliminación regresiva por paso (*stepwise backward elimination*) nos reveló que todos los efectos y sus interacciones contribuían a la bondad de ajustes del modelo, por lo que presentamos en la Tabla 2.7 la lista completa de los efectos fijos y sus coeficientes respectivos.

Tabla 2.7. Resultados del modelo regresión logística de efectos mixtos de la concordancia de género en español L3

	β	SE	z	Pr(> z)
(Intercept)	0.954	0.185	5.157	< 0.001
Grupo (FR)	-0.485	0.230	-2.107	0.035
Grupo (FR/EU)	-0.840	0.214	-3.926	< 0.001
Transparente	1.564	0.367	4.257	< 0.001
Congruente	0.571	0.463	1.234	0.217
Tipo (ADJ)	-0.286	0.149	-1.927	0.054
Género (FEM)	-3.982	0.465	-8.570	< 0.001
Género (FEM) × Tipo (ADJ)	-3.810	0.325	-11.736	< 0.001
Grupo (FR) × Transparente	-0.928	0.466	-1.991	0.046
Grupo (FR/EU) × Transparente	-1.451	0.436	-3.331	0.001
Grupo (FR) × Congruente	1.735	0.638	2.719	0.007
Grupo (FR/EU) × Congruente	0.578	0.584	0.991	0.322

Este modelo nos confirma la significatividad de los efectos globales observados en la Figura 2.3 y en las tasas medias de concordancia antes mencionadas, pero sobre todo nos informa sobre las interacciones entre los diferentes factores. El efecto de grupo es significativo para los tres

⁸⁶ Las demás interacciones impedían la convergencia del modelo. Además, según nuestro escrutinio visual de las gráficas de la Figura 2.3, los datos no justificaban la inclusión de otras interacciones en nuestro modelo.

grupos y los coeficientes van en el sentido de las tasas medias de concordancia observadas anteriormente, es decir, el grupo FR produce menos concordancia que el grupo EU y, a su vez, el grupo FR/EU produce significativamente menos concordancia que el grupo EU. En cuanto al factor género, se puede observar que el género femenino impacta de manera significativa y negativa sobre la concordancia de género, algo que también se ha comentado en trabajos anteriores (Keating, 2009; White et al., 2004) y que corresponde al llamado el efecto del masculino por defecto. Efectivamente, nuestros participantes parecen sobregeneralizar la concordancia con el masculino y evitar la concordancia en femenino. Además, el factor género interactúa de manera significativa con el tipo de concordancia, creando una condición poco favorable a la concordancia cuando el adjetivo tiene que concordar con un sustantivo femenino. El efecto tipo (ADJ) por sí solo es marginalmente significativo, ya que su interacción con el factor género (FEM) explica gran parte de la variación. A su vez, el factor de transparencia es significativo y positivo, es decir, la transparencia de la terminación favorece la concordancia. En cambio, el factor congruencia por sí solo no es significativo, pero, como detallaremos a continuación, su interacción con el factor grupo sí lo es.

Dos interacciones con el factor grupo destacan de este modelo: grupo \times congruencia y grupo \times transparencia. La primera confirma lo observado en la última gráfica de la Figura 2.3, es decir, que, en comparación con el nivel de referencia (grupo EU), el grupo FR realiza significativamente más la concordancia cuando el género del sustantivo en español es congruente con el género de su equivalente en francés. Como se puede observar en la última gráfica de la Figura 2.3, el grupo FR es el que muestra la mayor variación entre una condición y otra, pasando de una tasa de concordancia del 71,6 % en la condición congruente al 40,3 % en la condición incongruente. En cuanto a la interacción de los factores grupo \times transparencia, tenemos que tomar en cuenta el efecto global de transparencia (positivo y significativo) para interpretar la interacción que va en el sentido contrario (negativo). En comparación con el grupo EU, el efecto de la interacción entre la transparencia y los grupos FR y FR/EU es negativo, es decir, a nivel global, la transparencia favorece la concordancia, pero los grupos FR y FR/EU se benefician significativamente menos que el grupo EU de este efecto facilitador. Volveremos sobre estas interacciones en la discusión, pero, de manera provisional, podemos afirmar que los factores congruencia y transparencia afectan de manera distinta a los diferentes grupos, con el grupo EU mostrando una mayor sensibilidad a la transparencia y el grupo FR a la congruencia.

2.7.5 Español L3: perspectivas longitudinales

Para ver cómo evoluciona el dominio del fenómeno con el paso del tiempo en nuestros grupos experimentales, comparamos los datos de la primera recogida con las dos recogidas sucesivas que se llevaron a cabo después de 140 y 200 horas de enseñanza en español, respectivamente. Para comprobar si los efectos observados en la primera recogida de datos se mantenían en el tiempo, aplicamos el mismo modelo a los datos de las tres recogidas de manera separada y, para comparar la evolución entre recogidas, construimos un modelo similar al cual añadimos el factor tiempo (recogida 1, 2 o 3) y lo aplicamos al conjunto de datos de las tres recogidas.

Este último modelo nos permite comprobar la significatividad del aumento general de la tasa media de concordancia de los participantes, partiendo de un 55,89 % de concordancia en la primera recogida, para llegar a un 69,95 % en la segunda y a un 76,25 % en la última. Estos aumentos son significativos, tanto entre la primera y la segunda recogida ($\beta = 0.98816$, $SE = 0.08163$, $z = 12.105$, $p < .001$), como entre la segunda y la tercera ($\beta = 0.46963$, $SE = 0.08101$, $z = 5.797$, $p < .001$). La utilización de un modelo de regresión logística de efectos mixtos que no solo incluye el factor tiempo, sino todos los factores antes mencionados, nos permite comprobar que, aun considerando todos los demás factores significativos (grupo, género, transparencia, congruencia, grupo \times transparencia, grupo \times congruencia, género \times tipo), el factor tiempo es significativo.⁸⁷ Esta mejora de una recogida a otra responde a lo esperado y a lo que se documenta generalmente en adquisición de lenguas no nativas, es decir, con más horas de instrucción y exposición a la lengua, el nivel de los estudiantes mejora. Esto se aplica también a un fenómeno como el género gramatical, cuyas dificultades de aprendizaje se corroboran, pero no parecen ser insuperables.

Ahora, volviendo a los efectos observados en la primera recogida, podemos ver en las Figura 2.4 y 2.5 que estos parecen mantenerse. Juntos con las figuras, comentaremos los resultados de los modelos de regresiones logísticas de efectos mixtos que hemos aplicado a los datos de la segunda y la tercera recogida. Solo reportaremos los resultados más destacados y referimos al lector a los Anexos 7 y 8 para mayores detalles sobre cada modelo. La primera gráfica de la Figura 2.4

⁸⁷ Los detalles de este modelo se encuentran en el Anexo 6.

nos revela una evolución no lineal del uso del masculino en la condición femenina. Si bien en la primera recogida los participantes sobreutilizaban el masculino con los sustantivos femeninos (69,8 % de masculino de media en la condición femenina), a partir de la segunda recogida todos los grupos empiezan a usar el femenino en los contextos adecuados con más frecuencia (por encima del 50 %). A pesar de esa evolución, el efecto del género (femenino) sigue siendo muy significativo tanto en la segunda recogida ($\beta = -2.071, SE = 0.416 -4.976, p <.001$) como en la tercera ($\beta = -2.702, SE = 0.492 -5.489, p <.001$).

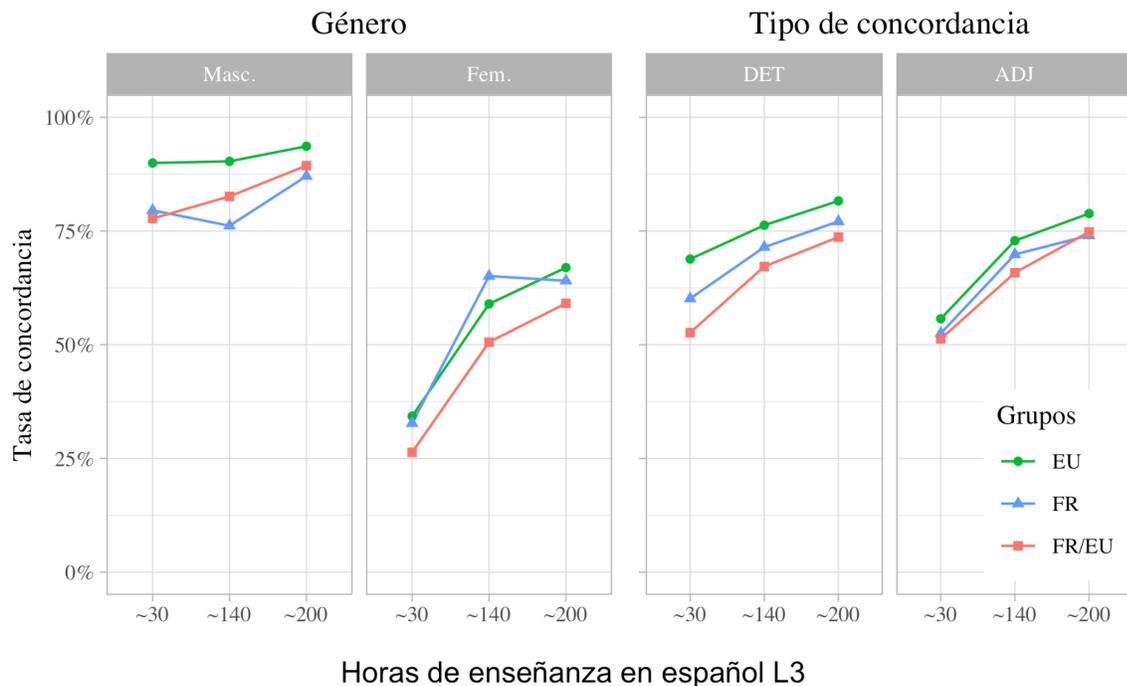


Figura 2.4. Evolución de la tasa media de concordancia por grupo según el género y el tipo de concordancia.

En cuanto al tipo de concordancia, el efecto que era solo marginalmente significativo en la primera recogida (72,83 % de concordancia con los determinantes, 67,25 % con los adjetivos) parece atenuarse con el tiempo, con una diferencia de alrededor del 2 % en la segunda recogida (70,97 % con los determinantes vs. 68,96 % con los adjetivos) y de 1 % en la última (76,79 % vs. 75,71 %). A su vez, la interacción género (FEM) \times tipo (ADJ) sigue siendo muy significativa en la segunda recogida ($\beta = -1.199, SE = 0.233, z = -5.155, p <.001$), pero no en la tercera ($\beta = -0.379, SE = 0.275, z = -1.376, p = .169$).

En cuanto a la transparencia, se puede observar en la gráfica de la izquierda de la Figura 2.5 que el grupo EU se mantiene por encima de los demás en la condición transparente y que la

interacción grupo \times transparencia sigue siendo significativa en la segunda recogida, aunque solo en el contraste EU - FR ($\beta = -1.412$, $SE = 0.594$, $z = -2.379$, $p = .017$). En la tercera recogida, la amplitud del efecto disminuye y la interacción de la transparencia con los grupos deja de ser significativa. Sin embargo, el efecto de transparencia por sí solo sigue siendo significativo tanto en la segunda recogida ($\beta = 1.907$, $SE = 0.491$, $z = 3.886$, $p < .001$) como en la tercera ($\beta = 3.024$, $SE = 0.661$, $z = 4.574$, $p < .001$). De hecho, la significatividad que aparecía en la interacción entre grupo \times transparencia se traslada al efecto de transparencia por sí solo, lo que apunta hacia una dificultad persistente con el género de los sustantivos con una terminación opaca en todos los grupos.

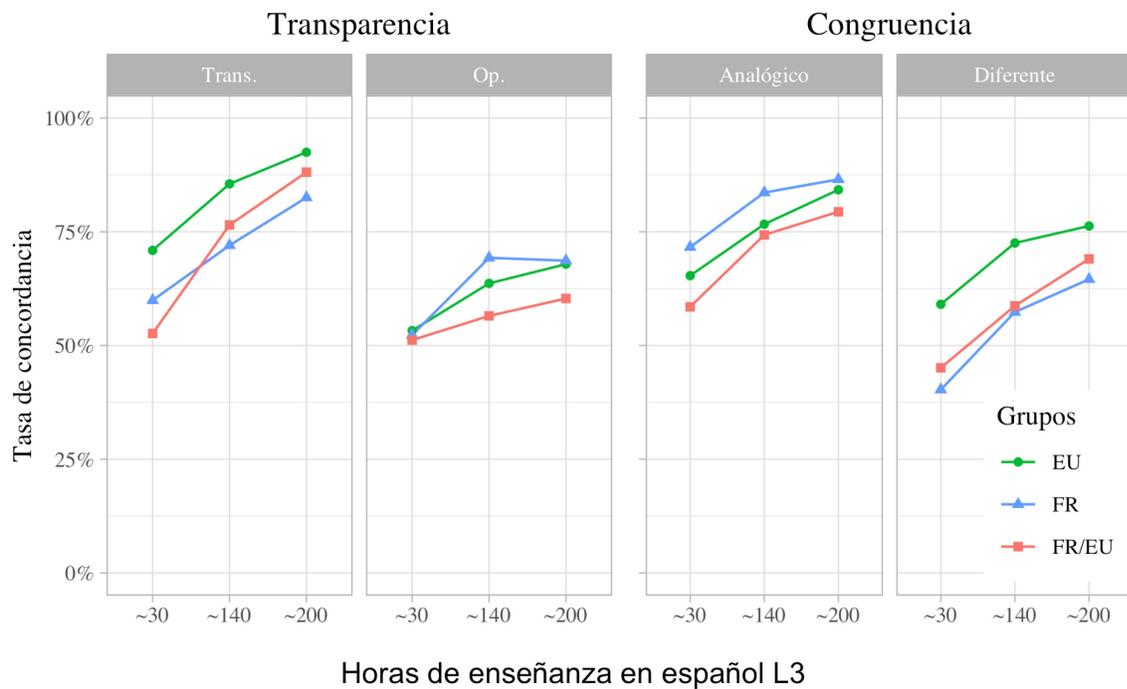


Figura 2.5. Evolución de la tasa media de concordancia por grupo según la transparencia y la congruencia

Con respecto a la congruencia, el grupo FR sigue teniendo mayor dificultad en la condición incongruente y la interacción grupo \times congruencia se mantiene de manera más duradera. En comparación con el nivel de referencia (EU), el efecto grupo (FR) \times congruente es significativo tanto en la segunda ($\beta = 1.660$, $SE = 0.526$, $z = 3.156$, $p = .002$) como en la tercera recogida ($\beta = 1.319$, $SE = 0.537$, $z = 2.457$, $p = .014$).

2.8 Discusión

Como vimos en la sección anterior, los resultados de la prueba de control en francés confirman, como era de esperar, la hipótesis 1, es decir, encontramos una alta tasa de aciertos en la concordancia de género entre el sustantivo y sus modificadores. Observamos una diferencia en función del tipo de concordancia (más aciertos con los determinantes) y del género (más aciertos en masculino), pero las variaciones no son muy prominentes, ya que todos nuestros participantes tienen un nivel nativo o casi nativo de esta lengua. El hecho de comprobar el nivel de dominio de la concordancia en francés para todos los participantes nos permitió considerar esta lengua como una posible fuente de influencia hacia la L3 en las demás hipótesis.

En cuanto a la hipótesis 2 sobre los factores observados en estudios anteriores de adquisición de L2, los resultados confirman que nuestros participantes muestran una mayor facilidad de concordancia con el determinante que con el adjetivo (2a), con los sustantivos masculinos que con los femeninos (2b) y con las terminaciones transparentes que con las opacas (2c). El factor interlingüístico del criterio analógico también parece jugar un papel facilitador (2d). De manera general, el proceso de adquisición del género en la L3 parece seguir las mismas pautas que en la L2. A continuación, indagamos en cada uno de estos factores, tras lo cual exploraremos de manera más específica la compleja influencia interlingüística propia del trío de lenguas de nuestro estudio y reflexionaremos sobre las implicaciones de nuestros resultados con respecto a los modelos de adquisición de L3.

Con una tasa de concordancia del 55,89 % en la primera recogida (y de apenas 30,28 % con los sustantivos femeninos), nuestros participantes claramente presentan dificultades con el rasgo de género en español como L3. A primera vista, resulta difícil discernir si se trata de un déficit del género inherente asignado a cada sustantivo o si se trata de un déficit en la operación de concordancia que depende del rasgo funcional de género (o una combinación de ambos). La tasa de aciertos en la condición masculina no aporta pistas sobre la operación de concordancia, ya que, como hemos visto, el género por defecto en español es el masculino. En ese sentido, puede interpretarse que la concordancia en masculino también refleja la ausencia de marcas de género. En cambio, el hecho de que la tasa de concordancia no sea nula en la condición femenina muestra que los participantes son capaces de realizar la concordancia en femenino en algunos casos. Sin embargo, esta operación de concordancia parece aplicarse a los determinantes antes que a los

adjetivos, ya que ciertos aprendientes ($n=16$) tienen una tasa de concordancia nula en la concordancia entre los sustantivos femeninos y los adjetivos en la primera recogida. A partir de la segunda recogida, la concordancia parece extenderse al adjetivo y los aprendientes ya no muestran ninguna diferencia significativa entre los dos tipos de concordancia (determinante vs adjetivo). Si bien algunos estudios demuestran que esta mayor dificultad de la concordancia con el adjetivo puede perdurar hasta los niveles intermedio y avanzado cuando se observa el procesamiento en tiempo real o cuando la concordancia es de larga distancia (Gillon Dowens et al., 2009; Keating, 2009), nuestra metodología no nos permite evaluar tales matices y consideramos que, dada la naturaleza de nuestra tarea, nuestros resultados son compatibles con los estudios anteriores de L2 (Bruhn de Garavito & White, 2002; White et al., 2004; entre otros).

En cuanto al factor del masculino por defecto, si bien los datos de la primera etapa apuntan a favor de una sobreutilización del masculino, como proponíamos en la hipótesis (2b), los datos longitudinales nos permitieron ver que esta estrategia perdía prominencia a medida que los participantes iban avanzando en su proceso de adquisición. Esta progresión nos lleva a afirmar que, si bien nuestros datos reflejan un uso del masculino por defecto, se trata de una estrategia inicial, ante la ausencia de otro tipo de recursos o estrategias (como el uso de indicios morfofonológicos).

Con respecto al papel de la terminación (hipótesis 2c) y al criterio analógico (hipótesis 2d), todos los grupos parecen beneficiarse en cierta medida de estos factores y el efecto se mantiene a lo largo de las tres recogidas. Sin embargo, las dos primeras lenguas de los participantes parecen modular el peso de estos factores, por lo que, a continuación, indagamos en el tema de la influencia interlingüística en más detalle.

Recordamos que, conforme a los modelos de adquisición de L3 que consideramos en esta tesis, era posible postular una influencia o bien a partir del euskera, o del francés. Según el *Typological Primacy Model* (TPM), el *Linguistic Proximity Model* (LPM) y el *Cumulative Enhancement Model* (CEM), la fuente de influencia privilegiada es el francés. En cuanto al *L2 Status Factor*, la influencia procede del francés solo si se trata de la L2 de los participantes. En el caso contrario, se predice una influencia del euskera. Con respecto al llamado *L1 Factor*, la fuente de influencia privilegiada es el francés solo si se trata de la L1 de los participantes y el euskera en el caso contrario. Finalmente, el *Competition Model* (CM) predice una influencia de la lengua dominante, por lo menos en las etapas iniciales. Considerando estas hipótesis de influencia,

detallamos los posibles efectos de cada lengua sobre el proceso de adquisición del español como L3. A continuación, reproducimos las predicciones para el género gramatical (G) en función las diferentes fuentes de influencia (EU o FR):

- G1-EU: El euskera no materializa el género gramatical, entonces, una influencia de esta lengua podría manifestarse con una tasa de concordancia más baja y con una sobreutilización del masculino, el género por defecto.
- G1-FR: Consideramos que los participantes que muestren una influencia del francés deberían beneficiarse de un efecto facilitador (mayor tasa de acierto de manera general), ya que los sistemas de género del español y del francés son similares.

Nuestros datos revelan que, a nivel global, el grupo EU obtiene los mejores resultados en la primera recogida, si bien la significatividad de este efecto de grupo desaparece en las recogidas posteriores. Considerando nuestra hipótesis (G1-FR), el grupo EU parece beneficiarse de una influencia facilitadora del francés. Por el contrario, los grupos FR y FR/EU obtienen resultados significativamente más bajos, lo que parece apuntar hacia una influencia inhibidora del euskera. Sin embargo, en cuanto al masculino por defecto, su interacción con el factor grupo no es significativa, por lo que no podemos concluir que los grupos FR y FR/EU sobreutilicen el masculino en mayor medida que el grupo EU. De hecho, el grupo que sobreutiliza más el masculino es el grupo EU, si bien esta interacción grupo \times género no resulta significativa. Considerando la falta de interacción entre los factores grupo y género por defecto, resulta difícil confirmar la hipótesis de la influencia del euskera (G1-EU), ya que la mayor tasa de errores de los grupos FR y FR/EU no parece proceder de la sobregeneralización del masculino en particular, como planteábamos en nuestra hipótesis. Entonces, concluimos que nuestros resultados no permiten desambiguar de manera clara entre las dos predicciones de la hipótesis G1.

Por otra parte, en lo relativo a la hipótesis G2 sobre los factores de transparencia de la terminación y del criterio analógico (congruencia), la fuente de influencia es más clara.

- G2-EU: Si bien predecimos un efecto facilitador de la terminación transparente para todos los grupos, predecimos un efecto de mayor amplitud para los participantes que muestren una influencia interlingüística a partir del euskera.
- G2-FR: La influencia del francés debería modular el efecto del criterio analógico, es decir, el género analógico debería facilitar la concordancia de manera más acentuada en los participantes que muestren una influencia del francés.

En cuanto a la transparencia de la terminación, el grupo EU es el que muestra los mejores resultados en la condición transparente (G2-EU), mientras que, con respecto al criterio analógico, el grupo FR muestra los mejores resultados en la condición congruente (G2-FR). A su vez, el grupo FR/EU se acerca al grupo FR en cuanto a la terminación y al grupo EU en cuanto al criterio analógico. Concretamente, no parece beneficiarse de ninguno de los dos efectos facilitadores, por lo que obtiene la tasa de aciertos más baja. En las dos subsecciones siguientes, nos concentramos sobre los resultados de los grupos EU y FR, ya que presentan patrones más definidos.

2.8.1 El papel de los indicios morfofonológicos

Como mencionamos, los resultados del grupo EU destacan particularmente con respecto al factor de la transparencia de la terminación. En la condición transparente, este grupo obtiene resultados significativamente más altos que los demás grupos (un 70,92 % frente a un 59,92 % para el grupo FR y un 52,64 % para el grupo FR/EU). Interpretamos estos resultados como la muestra de una mayor sensibilidad general a los indicios morfofonológicos motivada por las características específicas del euskera. Como indican Caffara et al. (2017) "[...] when bilinguals are used to speaking a gender-free agglutinative language (i.e. Basque) the lexical representation of gender is less stable and formal cues are more influential in agreement processing" (p. 14). Por un lado, el euskera es una lengua que no tiene género, lo que puede debilitar el rasgo del género en francés, la otra lengua hablada por nuestros participantes. Por otro lado, al ser una lengua aglutinante, una gran parte de la información morfosintáctica del euskera se presenta en los morfemas finales, por lo que los hablantes de esta lengua tienen que prestar mayor atención a las terminaciones de las palabras de manera general. Por consiguiente, podemos considerar que, como los participantes del grupo EU tienen una representación del género menos estable en francés, no se guían tanto como el grupo FR por el criterio analógico para asignar un género a los sustantivos en español, sino que recurren a la forma de la terminación.

Por otra parte, el efecto observado en ese grupo podría proceder también (o en parte) del hecho de que estos participantes han adquirido el francés como L2, es decir, en un contexto más formal, y que, durante su proceso de adquisición, han desarrollado cierta sensibilidad hacia el vínculo entre morfofonología y género en esa lengua. A la hora de adquirir otra lengua con las mismas categorías de género, en lugar de solamente usar el mismo género para los sustantivos equivalentes de una lengua a otra, los aprendientes del grupo EU parecen analizar las terminaciones

y atribuir un género al sustantivo en español basándose en los indicios morfofonológicos. En este caso, podríamos afirmar que hay cierta influencia a partir del francés, la L2, pero que no se materializa como una influencia interlingüística como tal, sino como un aprovechamiento de estrategias metalingüísticas entre los procesos de adquisición de L2 y L3. Se trataría, siguiendo a Falk et al. (2015), de un efecto de la instrucción en la L2.

En ambos casos, nuestros resultados apuntan al hecho de que los participantes del grupo EU (1) recurren menos al rasgo de género de la lengua más cercana tipológicamente y (2) recurren más a sus conocimientos metalingüísticos, ya procedan de la L1 (euskera) o de la L2 (francés).

2.8.2 El papel del criterio analógico

En cuanto al grupo FR, la principal fuente de influencia a la hora de asignar un género a un sustantivo en español parece ser el rasgo de género del sustantivo equivalente en francés. La diferencia entre las dos condiciones es bastante clara en la primera recogida, en la que este grupo obtiene la tasa más alta de concordancia en la condición congruente (un 71.6 %, en comparación con un 64.2 % para el grupo EU y un 59.3 % para el grupo FR/EU) y la más baja en la condición contraria (un 40.3 % en comparación con un 58,3 % para el grupo EU y un 45,8 % para el grupo FR/EU). Esta diferencia se mantiene a lo largo de las tres recogidas y sigue siendo significativa, pero con una amplitud menor.

Con respecto al criterio analógico, podemos establecer un vínculo entre lo que observamos en el marco de nuestro estudio y lo que se observó en estudios sobre la alternancia de código en los SD mixtos. Del mismo modo que observamos un efecto predominante del criterio analógico entre el francés y el español para los participantes que tenían el francés como lengua dominante, varios estudios sobre la alternancia de código reportan una influencia del criterio analógico solo cuando la lengua dominante tiene un sistema de género, como vimos en la sección 2.4.2. Para ilustrar el fenómeno, seguimos a Klassen y Liceras (2017) y aplicamos el Mecanismo de cotejamiento de los rasgos dobles de género (*Gender Double-Feature Valuation Mechanism*; Liceras, Fernández Fuertes, et al., 2008) a algunas respuestas de nuestros participantes del grupo FR. En el ejemplo (2.31), la analogía con el francés resulta provechosa porque el francés y el español tienen un género congruente. Además, la terminación opaca del sustantivo (-e) no proporciona ningún indicio potencial de género.

2.8.3 Influencia interlingüística y conocimientos metalingüísticos

En resumen, los grupos EU y FR muestran patrones diferentes en la concordancia de género en español L3. Por una parte, el grupo EU muestra una mayor sensibilidad a la terminación de los sustantivos, un efecto que se podría atribuir a unos conocimientos metalingüísticos precisos sobre la morfología. Sin embargo, no podemos desambiguar entre las dos posibles fuentes de influencia: el efecto de la instrucción de la L2 (francés) o el efecto de los procesos morfológicos de la L1 (eusquera) aplicados a la L3. En el caso del francés, estos resultados estarían en línea con el *L2 Status Factor*, pero si atribuimos esta influencia al euskera, estaríamos más bien hablando de un efecto del *L1 factor*. Por otra parte, la mayor sensibilidad del grupo FR al criterio analógico se podría atribuir a la proximidad entre el francés y el español, como sugieren el *Typological Primacy Model* y el *Linguistic Proximity Model*. Sin embargo, estos modelos no predicen ninguna diferencia en función del orden de adquisición o de la lengua dominante, pero en nuestros resultados sí se observan diferencias en ese sentido. Así pues, sugerimos que la mayor sensibilidad a la congruencia del grupo FR ocurre porque sus representaciones de género son más estables y más activas al contar ya con este rasgo en su L1 (Franceschina, 2005). Por lo tanto, esas representaciones son más difíciles de inhibir cuando se usa un sustantivo en español cuyo género contrasta con el de su equivalente en francés. En este sentido, el *Competition Model* parece predecir correctamente que la lengua dominante juega un papel principal en la interacción interlingüística.

En el próximo capítulo, exploramos la formación de los compuestos en la L3 de nuestros participantes con el objetivo de observar la potencial influencia interlingüística en otro fenómeno. Posteriormente, volveremos sobre los dos fenómenos en la discusión general (capítulo 4).

“[...] compounding is a field of study where intricate problems abound, numerous issues remain unresolved, and convincing solutions are generally not so easy to find” (Plag, 2003, p. 132)

3 La formación de los compuestos

Cada lengua favorece determinados procesos para formar nuevas palabras o expresiones, entre los cuales encontramos la derivación y la composición. Se considera generalmente que la derivación es más productiva que la composición en las lenguas romances, pero eso no impide que se haya publicado, a lo largo de los años, una cantidad considerable de trabajos relacionados con la composición en español. Más precisamente, varios trabajos en el campo de la adquisición de lenguas segundas (AL2) se han enfocado en los procesos de formación de unidades léxicas nuevas a partir de la unión de dos sustantivos (los llamados compuestos N+N, como “hombre lobo”) o de un verbo y un sustantivo (los llamados compuestos deverbales, como “limpiabotas”) (Lardiere & Schwartz, 1997; Liceras et al., 2002, 2004; Liceras & Díaz, 2000b; Liceras & Klassen, 2019; entre otros). Si bien existen otros tipos de compuestos, nos centraremos en estos dos tipos de compuestos nominales –llamados así porque la unidad resultante equivale a un sustantivo– ya que la literatura sobre la AL2 nos ofrece un punto de comparación para anclar nuestro estudio de AL3. Además, las tres lenguas de nuestro estudio presentan similitudes y diferencias con respecto a esos compuestos que nos permitirán comprobar nuestras hipótesis de influencia interlingüística.

La literatura sobre los compuestos en español tiene sus ejemplos clásicos como “bocacalle”, “hombre rana” o “abrelatas”, pero, de un autor a otro, la clasificación de esos mismos compuestos puede presentar complicaciones si entramos en los detalles taxonómicos.⁸⁸ Algunos autores consideran que los compuestos que llevan dos lexemas unidos sin espacio son los únicos compuestos propios o léxicos, mientras que otros incluyen en esa categoría también los compuestos

⁸⁸ El ejemplo más citado entre las referencias del presente capítulo es “abrelatas”, con 28 artículos, capítulos de libro o libros que lo mencionan como un ejemplo prototípico de los compuestos del español (Alcaraz Varó & Martínez Linares, 1997; Álvarez et al., 1997; Butt & Benjamin, 2011; García Mayo, 2006; Garza-González, 2014; Giurescu, 1975; González Alonso et al., 2016; Güemes et al., 2019; Gutiérrez Rodríguez, 2015; Kornfeld, 2009; Lardiere, 1995, 1998; Lardiere & Schwartz, 1997; Liceras & Klassen, 2019; Moyna, 2000, 2011; Orduña-López, 2019; Pérez-Leroux & Liceras, 2002; Pomerleau, 2000; Roca, 1989; Rossowová, 2007; Serrano Dolader, 2018; Tenorio Rebolledo, 2005; Tuggy, 2003; Varela, 2012; von Heusinger & Schwarze, 2013; Yang, 2016; Yoon, 2009).

en los que los lexemas están separados ortográficamente por un espacio, pero que muestran unidad semántica o nocional, como “hombre rana”.⁸⁹ Los primeros llaman compuestos sintagmáticos o impropios a expresiones como “hombre rana”, ya que las interpretan no como unidades léxicas, sino como sintagmas. A veces, se incluye también en los compuestos sintagmas nominales modificados por un sintagma preposicional, como “barco de vela”, mientras que otros autores los llaman compuestos prepositivos o sinápticos, o los excluyen simplemente porque se pueden formar mediante operaciones puramente sintácticas.

A pesar de los esfuerzos de los estudiosos por establecer una clasificación de los compuestos (Bauer, 2001; Scalise & Bisetto, 2011), los problemas terminológicos persisten y resulta, por lo tanto, imprescindible dedicar la primera parte del presente capítulo a la presentación de nociones preliminares que nos permitirán delimitar de manera clara los compuestos que entran en el marco de este estudio y la terminología que se usará a lo largo del capítulo. En la segunda parte del capítulo, describiremos cómo el fenómeno de la composición se manifiesta en las lenguas de nuestro estudio. El objetivo de la parte descriptiva consiste en comparar las tres lenguas en función de criterios que, si bien nos resultan útiles desde un punto de vista teórico, podrían no corresponder a las intuiciones de los hablantes, ya sea en su lengua nativa o no nativa. Dado que la literatura revisada en estas dos primeras partes procede de distintos marcos teóricos, trataremos de usar una terminología precisa, pero la más neutra posible en cuanto a la perspectiva, ya que la literatura revisada procede de distintos marcos teóricos. En la tercera parte, que concierne los análisis morfosintácticos y semánticos de los compuestos, presentaremos los aspectos clave de diferentes propuestas formales para poder compararlas y referiremos el lector a trabajos pertinentes para profundizar en cada análisis. En la sección 3.4, revisaremos los estudios anteriores sobre la adquisición de los compuestos en L1, 2L1, L2 y L3. En la sección 3.5, expondremos cómo la información presentada en las cuatro primeras partes del capítulo ha guiado el diseño de nuestro

⁸⁹ Aquí preferimos el término “lexema” a “palabra”, por una parte, porque el lexema hace referencia a una unidad mínima que puede ser o no flexionada, y, por otra, porque la palabra suele ir separada por unos espacios ortográficos, lo que no siempre se aplica a los componentes dentro de un compuesto. En la introducción a su libro *The notion of Word*, Di Sciullo & Williams (1987) enumeran tres “ideas” que suelen corresponder con la noción de “palabra”: un objeto morfológico (lo que entendemos aquí por lexema), un átomo sintáctico (que podría ser formado por varios lexemas) o un elemento memorizado (*listeme*, este tipo de “palabra” corresponder a un lexema irregular a cualquiera expresión fija).

estudio. Tras haber establecido estas bases, presentaremos nuestra metodología en la sección 3.6, expondremos nuestros resultados en la sección 3.7 y concluiremos con una discusión sobre estos resultados en la sección 3.8.

3.1 Nociones preliminares

Como subrayan Guevara y Scalise (2009, pp. 106-107), presentar una definición universal del compuesto o del proceso de composición propiamente dicho resulta problemático, ya que existen varios tipos de compuestos y su formación varía de un idioma a otro. Además, si bien algunos autores explican la formación de los compuestos a partir de procesos estrictamente morfológicos (Fradin, 2009; Guevara & Scalise, 2009; entre otros), otros hablan más bien de procesos que implican operaciones sintácticas (Barbaud, 1997; Roeper et al., 2002; entre otros). Se trata, dependiendo del punto de vista que se adopte y del tipo de compuesto considerado, de un fenómeno a caballo entre la morfología y la sintaxis.

En el presente trabajo, definiremos la composición de manera general como un proceso que consiste en combinar dos (o más) lexemas libres para obtener una nueva unidad con significado propio. Por ejemplo, a partir de los lexemas “perro” y “policía” se puede formar el compuesto “perro policía”. Con esta definición incluimos las estructuras presentadas en (3.1),⁹⁰ pero excluimos los llamados compuestos etimológicos, cuyos lexemas a veces no son claramente identificables, como en (3.2), y los llamados compuestos neoclásicos formados a partir de elementos de origen griego que no pueden aparecer de forma libre en el español actual, como en (3.3).

- (3.1) a. sustantivo (N) + sustantivo (N): “perro policía”, “telaraña”
b. verbo (V) + sustantivo (N): “espantapájaros”, “chupamedias”
c. preposición (P) + sustantivo (N): “contraindicación”, “sinsentido”
d. sustantivo (N) + preposición “de” + sustantivo (N): “barco de vapor” o “barco de vela”
e. sustantivo (N) + adjetivo (A): “carapintada”, “pelirrojo”
f. adjetivo (A) + sustantivo (N): “altavoz”, “malasangre”
g. adjetivo (A) + adjetivo (A): “sordomudo”, “blanquiceleste”
- (3.2) a. carcomer (de carne + comer) (Moyna, 2011, p. 35)
b. hilván (de hilo + vano)

⁹⁰ Los detalles de los procesos de amalgamación de “tela” y “araña”, o del uso de la vocal de enlace -i- en la unión “pelo” y “rojo”, o “blanco” y “celeste” quedan fuera del marco del presente trabajo. La lista de las estructuras y la mayoría de los ejemplos que las ilustran proceden de Kornfeld (2009).

- (3.3) a. psicología (de psico + logía) (Moyna, 2011, p. 42)
b. agricultura (de agri + cultura)

En términos generales, esta definición resulta satisfactoria, pero la matizaremos a lo largo de las próximas secciones y añadiremos detalles en función de las características de las lenguas de nuestro estudio. A partir de aquí, siguiendo las siglas introducidas en los ejemplos de (3.1), se usará la fórmula X+X para referirse a los componentes de un compuesto y, cuando sea relevante hacerlo, añadiremos corchetes y una sigla para representar la categoría resultante del compuesto.⁹¹ Así, la estructura [N+N]_N representaría un compuesto formado a partir de dos sustantivos cuyo resultado equivale a un nuevo sustantivo, por ejemplo, “perro policía”. A continuación, proponemos una serie de nociones preliminares que nos servirán de base para abordar la descripción de las lenguas del estudio (sección 3.2) y las propuestas de análisis teóricos (sección 3.3).

3.1.1 Núcleo, endocentricidad y exocentricidad

Para entender el fenómeno de la composición, la noción de núcleo es fundamental, ya que de ella dependen las nociones de endocentricidad y exocentricidad, que son omnipresentes en la literatura sobre los compuestos y que presentaremos a continuación.

La noción de núcleo procede originalmente del campo de la sintaxis y se refiere al elemento que domina un sintagma, pero desde los años ochenta también ha jugado un papel central en el ámbito de la morfología (Corbett et al., 1993; Fábregas & Masini, 2015; Zwicky, 1985). De manera general, se considera que el núcleo es el elemento que se proyecta y que sus propiedades se transfieren a la proyección que lo domina, es decir, el sintagma en el caso de la sintaxis, la palabra en el caso de la morfología.⁹² Refiriéndose específicamente a los compuestos, Scalise y Fábregas (2010) indican que, en una situación ideal, el núcleo debería transferir al compuesto todos sus

⁹¹ En esta representación, no hacemos distinción entre los compuestos unidos sin espacio, con guion o separados por un espacio. Si bien somos conscientes de que algunas obras de referencias, como la NGLE, representan la distinción entre los compuestos separados por un espacio o un guion como (X+X) y los fusionados como (X-X), esta diferencia no es relevante en el marco de nuestro estudio, como expondremos en la sección 3.1.6. En algunas ocasiones, también usaremos números para indicar la posición de un elemento dentro del compuesto (N1, N2)

⁹² Dado que algunos autores analizan los compuestos como unidades sintagmáticas formadas esencialmente a partir de la sintaxis, mientras que otros los analizan como productos meramente morfológicos, intentamos usar términos generales en esta sección. Entraremos en los términos propios de cada marco teórico en la sección 3.3.

rasgos categoriales, semánticos y morfológicos (el *locus inflectionis*), pero que no siempre ocurre así.

Con respecto a los rasgos categoriales, si tomamos el ejemplo del N+N “perro policía”, observamos que tanto la categoría de “perro” como la de “policía” (sustantivo) se corresponden con la categoría del compuesto resultante, por lo que los dos elementos podrían ser candidatos potenciales para ejercer de núcleo. En cuanto a la información semántica, el elemento “perro” es más prominente, ya que un “perro policía” es un tipo de “perro”, es decir que “perro” es el hiperónimo de “perro policía”. En (3.4) y (3.5) figuran dos pruebas simétricas usadas en Arnaud y Renner (2014), tomadas de Riegel (1988b) y Cruse (1986), y que aplicamos a “perro policía” para destacar el elemento que aporta la mayor información semántica al compuesto, el núcleo.

(3.4) Un N1N2 es un N1.

Un *perro policía* es un *perro*.

(3.5) Un N1N2 es un N2.

*Un *perro policía* es un *policía*.

Por otro lado, al cambiar el género del N1 (“una perra policía”), al flexionarlo en número (“dos perros policía”), o al derivarlo con un diminutivo (“un perrito policía”), se puede constatar también que los rasgos morfológicos de flexión se concentran en “perro” y se aplican al compuesto como conjunto.⁹³ Por lo tanto, podemos concluir que, en el caso de “perro policía”, los tres tipos de rasgos se proyectan del núcleo al compuesto entero. Sin embargo, identificar un núcleo que corresponda a los tres criterios de Scalise y Fábregas (2010) resulta más difícil en compuestos como los que aparecen en los ejemplos (3.6) a (3.8).

(3.6) telaraña

(3.7) cabeza rapada

(3.8) limpiabotas

En el ejemplo (3.6), que presenta también la estructura N+N, tanto “tela” como “araña” corresponden a la categoría del compuesto resultante (sustantivo). Con respecto a las pruebas (3.4) y (3.5), “tela” aparece como el elemento que aporta la mayor información semántica. Sin embargo, en cuanto a los rasgos morfológicos de flexión, resulta imposible flexionar “tela” en número (“dos *telasaraña”) y si queremos pluralizar el compuesto, no tenemos otra opción que colocar la -s al

⁹³ En la sección 3.4.3, al hablar de un trabajo de Klassen y Licerias (2019) sobre la adquisición de los compuestos en L2, aportaremos más detalles sobre los efectos de la flexión en los compuestos N+N.

final de este (“dos telarañas”). A pesar de esto, se suele considerar que “tela” es el núcleo de “telaraña”, por lo que el núcleo no parece ser siempre el lugar de flexión en los compuestos lexicalizados del español.⁹⁴

El ejemplo (3.7), un N+A, muestra otras particularidades. Tomando en cuenta el criterio categorial, solo podemos considerar “cabeza”, un sustantivo, como el núcleo del compuesto. Sin embargo, un “cabeza rapada” no es un tipo de “cabeza” ni un tipo de “rapada”, sino “una persona que tiene la cabeza rapada”, por lo que el compuesto no parece contener la información semántica que denota el tipo de referente al que se refiere el compuesto. En cuanto a los rasgos morfológicos, tanto “cabeza” como “rapada” se pueden flexionar en número (“dos cabezas rapadas”), pero en cuanto a la flexión de género, “rapada” siempre parece concordar con el sustantivo “cabeza” y no con el referente del compuesto, ya que se puede decir “una cabeza rapada” o “un cabeza rapada”, según nos referimos a una mujer o a un hombre, respectivamente. Es decir que, a nivel sintáctico, “cabeza” parece ser el núcleo de la construcción “cabeza rapada”, ya que impone su rasgo ϕ [+femenino] al adjetivo “rapada”; sin embargo, la ausencia de concordancia con el determinante nos revela que no parece ser el núcleo del compuesto.

El ejemplo (3.8), que presenta la estructura V+N, tiene otras particularidades distintas. Según el criterio de la categoría, el núcleo tendría que ser “botas”, al ser el un sustantivo, que es, a su vez, la categoría del compuesto. Sin embargo, ni “limpia”, ni “botas” cumplen con las pruebas semánticas. En cuanto a la flexión, los dos elementos se mantienen inalterados, independientemente del género y número del referente (“un limpiabotas”, “una limpiabotas”, “dos limpiabotas”).

Considerando los últimos cuatro ejemplos, se puede notar que los tres criterios –categorial, semántico y morfológico– no siempre están reunidos en un mismo elemento. Basándose en el análisis de 3000 compuestos su base de datos MorBoComp, que recoge compuestos de 21 lenguas, Scalise y Fábregas (2010) observan que cuando se puede identificar un núcleo en función del criterio semántico, este suelen coincidir con el criterio categorial lo que no es siempre el caso para

⁹⁴ Según Dressler (2006), los compuestos se vuelven cada vez menos transparentes a medida que avanzan en el progreso de lexicalización. Si bien se suele aludir a la transparencia en términos semánticos, también parece reflejarse en la forma de los compuestos altamente lexicalizados como “telaraña”. La opacidad en la ortografía (unión sin espacio) provoca mayor opacidad en la morfología (imposibilidad de identificar el núcleo mediante la flexión).

el criterio morfológico. Esta observación coincide con varios autores que consideran que el núcleo es, ante todo, el elemento que conlleva la mayor carga semántica y que le da su categoría al compuesto (Benczes, 2006; Bloomfield, 1933/1973; Booij, 2009; Kornfeld, 2009; Varela, 2012; entre otros).

En función de esta noción de núcleo, los compuestos se dividen en dos categorías: los que llevan un núcleo explícito, como “perro policía” y “telaraña”, se denominan endocéntricos y los que llevan un núcleo implícito, como “cabeza rapada” y “limpiabotas”, se consideran como exocéntricos. Cabe mencionar que Varela (1989, 2012) considera que los compuestos V+N son endocéntricos y que el elemento verbal (“limpia”) es en realidad un verbo sustantivado (“el limpia”), apoyándose en el hecho de que otros verbos sustantivados se pueden encontrar como sustantivos independientes (“guarda”, “caza”, “okupa”).⁹⁵

3.1.2 Subordinación y coordinación

Además de la jerarquía que se establece entre los elementos de un compuesto según su carácter nuclear o no nuclear, paralelamente, existe otra relación semántica y gramatical que puede ser de subordinación o coordinación. En los ejemplos presentados en la subsección anterior, siempre aparece una relación de subordinación en el sentido de que el segundo elemento está subordinado al primero. Esta relación puede ser atributiva como en los casos de “perro policía” o “cabeza rapada”, es decir, una relación modificado-modificador similar a la que existe entre un

⁹⁵ Los ejemplos de verbos sustantivados presentados en Varela (2012) presentan cierta incertidumbre en cuanto a su origen. Si bien Varela dice que “el limpia” se refiere a una persona que limpia, el DLE (RAE, 2014) lo presenta como un acortamiento de “limpiabotas”, por lo que se refiere más bien a “una persona que limpia las botas”, específicamente. Por su parte, el sustantivo “guarda” procedería del germánico “warda” que significa “acto de buscar con la vista” (RAE, 2014), por lo que su origen verbal tampoco está claro. Con respecto a “okupa”, el DLE lo presenta como un acortamiento de “ocupante” y no de “ocupar”. Además, Varela no precisa en qué compuesto deverbal se puede encontrar “okupa”, por lo que no queda claro cómo esta palabra respalda su análisis. En cuanto a “caza”, este también tiene una definición específica (“avión de caza” y no “caza-X”), por lo que tampoco queda claro su origen verbal. Además, si bien “caza” se encuentra en compuestos deverbales como “cazabobos” y “cazACLAVOS”, la clasificación de compuestos como “cazabombardero” no es tan obvia. Se podría analizar como un compuesto coordinativo formado por dos sustantivos ya que “bombardero” no es el complemento de “caza” sino una función adicional del avión que a la vez actúa como avión de caza y avión bombardero. Volveremos sobre los detalles de la propuesta de Varela en la sección 3.3.2.

nombre y un adjetivo en un sintagma nominal. A su vez, la relación puede ser de tipo predicado-argumento como en el caso de “limpiabotas”, donde “botas” es un argumento del verbo “limpiar”. Otra relación posible en los N+N es la “no atributiva” (RAE & ASALE, 2010a, pp. 196-197) o también llamada “relacional” (Arnaud, 2016) que es, como su nombre implica, una categoría que abarca una variedad de relaciones como las que existen en “telaraña” o “coche cama”.⁹⁶

En cambio, en otros compuestos, como en los ejemplos (3.9), no existe subordinación sino coordinación. Los elementos de esos compuestos están en un mismo nivel, de modo comparable a dos sintagmas coordinados por una conjunción coordinante, pero dentro de una misma unidad léxica. Para comprobar su jerarquía, siguiendo a Renner y Fernández-Domínguez (2011), se puede usar las pruebas que reproducimos en (3.10) y (3.11).

- (3.9) a. lavadora-secadora (Renner & Fernández-Domínguez, 2011, pp. 878-879)
 b. compositor-arreglista
- (3.10) X+Y también podría decirse Y+X
Lavadora-secadora también podría decirse *secadora-lavadora*
- (3.11) Un(a) X+Y es un(a) X que también es un(a) Y
 Una *lavadora-secadora* es una *lavadora* que también es una *secadora*
- (3.12) **Perro policía* también podría decirse *policía perro*
 *Un *perro policía* es un *perro* que también es un *policía*

Los compuestos coordinativos, también llamados co-compuestos o *dvandva*,⁹⁷ son productivos en español en las estructuras A+A, A+i+A y, en menor medida, en los N+N. Renner y Fernández-Domínguez (2011) también proponen pruebas para los compuestos coordinativos adjetivales, ya que las pruebas anteriores solo se aplican a los compuestos nominales. Las dos pruebas siguientes permiten dividir esos compuestos entre los aditivos (3.13) y los híbridos (3.14):

- (3.13) Compuestos coordinativos aditivos
 a. anchicorto, rojinegro, sordomudo, fisicoquímico, político-económico
 b. Ser X+Y es ser X y Y.
 c. Ser anchicorto es ser ancho y corto; ser rojinegro es ser rojo y negro; etc.
- (3.14) Compuestos coordinativos híbridos
 a. agridulce, verdiazul, claroscuro
 b. Ser X+Y es estar a medio camino entre X y Y.
 c. Ser agridulce es estar a medio camino entre agrio y dulce; etc.

⁹⁶ Volveremos sobre la naturaleza precisa de esas relaciones en la descripción del fenómeno en cada lengua del estudio.

⁹⁷ El término *dvandva* procede de los estudios sobre los compuestos en sánscrito y, por tradición, se sigue mencionando en trabajos recientes sobre la clasificación de los compuestos (Scalise & Bisetto, 2011; Bauer, 2001; entre otros).

Otra vez, estas pruebas ilustran que los dos elementos de los compuestos coordinativos se ubican en un mismo nivel. Así, es posible considerar que esos compuestos tienen dos núcleos, o un solo núcleo conformado por los dos elementos o, como se suelen clasificar más a menudo, que no tienen un núcleo implícito y que son exocéntricos.

Hasta aquí hemos clasificado los compuestos según su endo-/exocentricidad y coordinación/subordinación, siguiendo la taxonomía de la NGLE (RAE & ASALE, 2010), pero a modo de comparación, comentamos brevemente otras clasificaciones. Bisetto y Scalise (2005) recopilan una lista de nueve propuestas distintas que van desde Bloomfield (1933/1973) a Bauer (2001). La mayoría de las clasificaciones presentadas en esta revisión de la literatura se basan en la presencia o no de un núcleo (endo-/exocentricidad) y en las relaciones semánticas y gramaticales entre los componentes (sintética, apositiva, copulativa, coordinativa, subordinativa, aditiva, posesiva, determinativa, etc.). Tomando en cuenta estas clasificaciones, Bisetto y Scalise elaboran su propia taxonomía, que se basa en tres grandes categorías: subordinativo, atributivo y coordinativo. A su vez, cada una de estas categorías se subdivide en endocéntrico y exocéntrico. En una versión actualizada para abarcar nuevos datos de lenguas no consideradas en su taxonomía de 2005 y alcanzar un mayor nivel de precisión, Scalise y Bisetto (2011) reestructuran y añaden un nivel intermedio a su jerarquía.⁹⁸ Si bien tal nivel de granularidad nos parece pertinente en una taxonomía de alcance universal, en el presente trabajo nos concentramos en las oposiciones binarias coordinativo/subordinativo y endocéntrico/exocéntrico, que por otra parte, representa la terminología más utilizada en los estudios sobre el tema.

3.1.3 Direccionalidad

La direccionalidad (*headedness* o *directionality* en inglés) es otra noción esencial para describir los compuestos. Se trata de una propiedad que varía de un idioma a otro y que parece estar relacionada, según observó Beard (1995, 1996), con el orden de los elementos dentro del sintagma nominal, por lo menos en los compuestos nominales endocéntricos. Así, esta variabilidad

⁹⁸ En la nueva clasificación, la categoría “subordinativo” (SUB) se subdivide en *ground* y *verbal-nexus*, la categoría “atributivo” (rebautizada “ATAP”) se subdivide en “atributivo” y “apositivo”. Cada una de estas nuevas subcategorías se subdivide a su vez en “endocéntrico” y “exocéntrico”. La categoría “coordinativo” (COORD) se subdivide directamente en “endocéntrico” y “exocéntrico”, sin nivel intermedio (Scalise & Bisetto, 2011, p. 50).

en el orden de los constituyentes se pone de manifiesto si comparamos lenguas como el inglés, cuyo orden canónico es [adjetivo + nombre], con lenguas como el español, que privilegia el orden [nombre + adjetivo], como ilustra el ejemplo (3.15). De manera paralela, encontramos el orden [modificador + núcleo] para los compuestos [N+N] del inglés, y el orden [núcleo + modificador] para los equivalentes en español, como se puede observar en el ejemplo (3.16). Con el fin de formalizar la aparente relación entre la direccionalidad de los compuestos y la direccionalidad de los sintagmas nominales, Beard (1996) formuló la hipótesis que reproducimos en (3.17).

(3.15) *red wagon*
 rojo vagón
 “vagón rojo”

(3.16) *police dog*
 policía perro
 “perro policía”

(3.17) *Base-Rule Ordering*
 “The subordinate constituent of a compound N assumes a position identical with the default position of a single, unmodified, lexical adjective in a NP.” (Beard, 1996, p. 4)

Para comprobar esta relación, Beard comparó 48 lenguas de una variedad de familias y llegó a la conclusión de que, salvo las excepciones encontradas en algunas lenguas norteamericanas (Navajo, Koasati, Kiowa, Dakota, Nahuatl), la relación entre las dos estructuras parecía robusta.⁹⁹ Bauer (2001) realizó un trabajo similar con una selección de 36 lenguas (algunas coinciden con las del estudio de Beard), y encontró una relación menos clara, con 10 lenguas de tipo [nombre + adjetivo] / [núcleo + modificador] y 9 lenguas de tipo [adjetivo + nombre] / [modificador + núcleo].¹⁰⁰ En su muestra, el tipo de contraejemplo más frecuente (11 lenguas) corresponde a la

⁹⁹ Beard enumera las 48 lenguas: “Amharic, Archi, Bashkir, Basque, Chukchee, Dakota, Dutch, English, Fijian, Finnish, French, German, Gilyak, Hausa, Hebrew, Hungarian, Indonesian, Italian, Kiowa, Koasati, Koryak, Latin, Lezghian, Lingala, Nahuatl (Huasteca, North Pueblo, and Tetelcingo), Navajo, Pashto, Polish, Quechua, Russian, Samoan, Sanskrit, Serbo-Croatian, Somali, Spanish, Swahili, Tagalog, Tamil, Thai, Turkana, Turkish, Veps, Vietnamese, Woleaian, Yakutsk, Yimas, Yoruba, and Yupik” (1996, p. 27).

¹⁰⁰ Bauer enumera las 36 lenguas examinadas según su área geográfica de origen: Africa (Ewe, Hebreo, Kanuri, Tswana, Turkana, Yoruba), Eurasia (Abkhaz, Euskera, Chukchee, Danés, Finés, Tamil), Sudeste asiático/Oceanía (Cantonés, Camboiano, maorí, tai, toba-batak, vietnamita), Australia y Nueva Guinea (Arabana-Wangkangurru, Kobon, Mara, siroi, waskia, yimas), América del Norte (dakota, kiowa, takelma, tumpisa, shoshone, tuztujil, oestegroenlandés) y América del Sur (cayuvava, guaraní, hixkaryana-imbabura, quechua, piraha) (2001, p. 696)

combinación [nombre + adjetivo] / [modificador + núcleo], el mismo orden que los contraejemplos de Beard. Comentando estos dos estudios, Snyder (2016a, p. 110) emite cierta reserva en cuanto a la metodología de Bauer (débil representatividad de las fuentes estudiadas, principalmente gramáticas descriptivas; inclusión demasiado permisiva de compuestos que posiblemente no lo eran) y prefiere seguir los resultados de Beard (1996). Cabe emitir cierta reserva también en cuanto a los datos presentados en Beard (1996) ya que, si bien el autor incluyó el euskera dentro de su muestra, no presentó esa lengua como una excepción a su hipótesis. Sin embargo, los ejemplos (3.18) y (3.19) muestran claramente que el orden canónico en euskera es de [nombre + adjetivo] (de Rijk, 2008, p. 18) como en español, pero los compuestos [N+N] se forman con la direccionalidad inversa, es decir, [modificador + núcleo] (de Rijk, 2008, p. 867), como en inglés.

(3.18) *bagoi gorri*

vagón rojo

“vagón rojo”

(3.19) *polizia-txakur*

policía perro

“perro policía”

Ampliando la muestra de lenguas, el *World Atlas of Language Structures* (Dryer, 2013) nos revela que el orden [nombre + adjetivo] es el más frecuente en las lenguas del mundo con un total de 878 lenguas que privilegian dicho orden frente a 378 que privilegian el orden inverso y 110 lenguas que no muestran una clara preferencia por un orden u otro. Desafortunadamente, el *WALS* no proporciona ningún dato sobre la direccionalidad de los compuestos, por lo que no podemos comprobar la relación entre las dos estructuras en esta muestra de lenguas más amplia. Sin embargo, si consideramos que varios autores (Dressler, 2006; Guevara & Scalise, 2009; entre otros) indican que el orden más frecuente en los compuestos a través las lenguas del mundo es el de [modificador + núcleo], observamos una aparente asimetría entre el orden dominante dentro del sintagma nominal y la direccionalidad de los compuestos nominales, la cual parece contradecir la *Base-Rule Ordering* de Beard (1996). En vista de todo esto, más que una regla, parece existir cierta tendencia entre la direccionalidad del sintagma nominal y la de los compuestos nominales, pero nos parece relevante destacar que esta tendencia no se aplica a una de las lenguas de nuestro estudio, el euskera.¹⁰¹

¹⁰¹ Volveremos sobre la direccionalidad de los compuestos en cada lengua en la sección 3.2.

3.1.4 Atomicidad

Los compuestos se diferencian de los simples sintagmas por su llamada atomicidad, es decir, no permiten la inserción, el movimiento y las relaciones anafóricas (Scalise, 1992, p. 195). Di Sciullo y Williams (1987) los consideran como “palabras sintácticas” (*syntactic words*) para aludir a la vez a su aparente formación sintáctica, pero también a su atomicidad propia de las palabras. Por ejemplo, no es posible insertar el adverbio “muy” dentro de “perro policía” para modificar “policía”, como en el ejemplo (3.20), mientras que sí sería posible insertar el mismo adverbio entre un sustantivo y un adjetivo que lo complementa, como en (3.21).

(3.20) *un perro muy policía

(3.21) un perro (muy) amable

En cuanto al movimiento, es imposible mover solo una parte del compuesto para formular una pregunta ya que el compuesto representa una isla anafórica. Esto significa que no es posible hacer referencia solo a una de sus partes, como en el ejemplo (3.22). Tampoco se puede calificar solo una parte del compuesto, como en el ejemplo (3.23).

(3.22) *el [[limpia[botas_i]] las_i limpia muy bien

(3.23) *un [[limpia[botas_i]] negras_i

Como se puede observar en los ejemplos anteriores, los compuestos parecen tener una estructura interna hermética a la sintaxis, lo que lleva a ciertos autores a considerarlos desde una perspectiva puramente morfológica, como lo veremos en algunos análisis teóricos de la sección 3.3. A pesar de la aparente impermeabilidad a la sintaxis, si bien no se puede insertar palabras dentro de un compuesto, los compuestos N+N con núcleo inicial sí permiten la afijación de morfemas flexivos sobre su núcleo, como observamos en los ejemplos (3.24) y (3.25).

(3.24) una perra policía

(3.25) dos perrosos policía

(3.26) dos perrosos policías

Si bien el ejemplo (3.25) solo muestra la marca del plural en el núcleo, el español también permite la concordancia en número con el elemento no nuclear como en (3.26), aunque existe una clara preferencia para marcar solo el plural en el núcleo (Guevara & Scalise, 2009, p. 124). Volveremos sobre esta discusión de manera más detallada a la hora de describir el fenómeno en cada lengua del estudio (sección 3.2).

3.1.5 Compuestos sintagmáticos y compuestos propios

Considerando solo el criterio de atomicidad, tendríamos que incluir dentro de los compuestos también sintagmas nominales que contienen una preposición intercalada, como en los ejemplos (3.27a) y (3.28a). Las variantes (3.27b-c) y (3.28b-c) muestran que no es posible insertar palabras dentro de esos sintagmas, calificar un elemento no nuclear o hacer referencia a una parte del sintagma sin cambiar su significado.

- (3.27) a. molino de viento
- b. *molino de mucho viento
- c. *molino de vientos septentrionales
- (3.28) a. dulce de leche
- b. *dulce de buena leche
- c. *dulce de leche_i [que produce mi vaca]_i

Sin embargo, autores como Rainer y Varela (1992) consideran esos ejemplos como compuestos impropios, ya que su estructura interna sigue las reglas de los sintagmas nominales del español. Benveniste considera esas estructuras a la vez como diferentes del compuesto y de los sintagmas y utiliza un término específico para referirse a ellas: “sinapsias” (1966/1974, p. 172). En el marco del presente trabajo, nos referiremos a esos casos como compuestos sintagmáticos preposicionales (o simplemente [N+de+N] para los que presentan la preposición “de”, o [N+prep+N] para referirnos a estos compuestos de manera más general), en contraste con los propios, o N+N, los cuales consideraremos que no pueden llevar una preposición intercalada.

3.1.6 Ortografía y patrones de acentuación

La ortografía de los compuestos presenta cierta variación en torno a dos aspectos: la unión y la pluralización de sus componentes. En cuanto a la unión de los componentes, ciertos compuestos se escriben con un guion (p. ej., “poeta-pintor”), otros con un espacio (p. ej., “perro policía”) y otros en una sola palabra (p. ej., “telaraña”). En lo relativo a la pluralización, en algunos casos se permite la pluralización de ambos componentes (p. ej., “dos perros policías”), otras veces solo uno (p. ej., “dos telarañas”), y, en ocasiones, la marca del plural solo tiene un valor genérico, por lo que el compuesto puede llevar la terminación -s sin concordar con otros elementos externos (p. ej., “un limpiabotas”). Dado que cada lengua tiene sus particularidades a este respecto, la descripción detallada sobre cada lengua de nuestro estudio se presentará en la sección 3.2.

Por su parte, los patrones de acentuación de los compuestos han recibido mucha atención en lenguas como el inglés, ya que la composición suele modificar la acentuación original de los componentes de los compuestos y ciertos autores han postulado reglas específicas –con sus respectivas excepciones– para la acentuación de los compuestos (Plag, 2006).¹⁰² En cambio, en español, los patrones han sido menos estudiados y los pocos estudios disponibles indican dos posibilidades: la opcional pérdida de acento del primer componente, como vemos en el ejemplo (3.29), o la conservación de los acentos en ambos componentes, como en los ejemplos (3.29) y (3.30). Como indica Moyna (2011, p. 81), la primera opción suele indicar una mayor lexicalización del compuesto en cuestión.¹⁰³

- (3.29) a. camposanto /kan.po.sán.to/
b. campo santo /kám.po.sán.to/
(3.30) hombre lobo /hóm.bre.ló.bo/

Como se trata de un área poco documentada en las tres lenguas del estudio, resultaría difícil establecer hipótesis de influencia interlingüística, por lo que no tomamos en cuenta la acentuación de los compuestos en el diseño de nuestro estudio.

3.1.7 Productividad, creatividad y recursividad

La productividad –o creatividad como Snyder (2016b) prefiere llamarla– de un fenómeno lingüístico como la composición consiste en la posibilidad de crear nuevas palabras o expresiones en una lengua dada de manera libre, sin que esa expresión haya sido acuñada antes. En el caso de los compuestos N+N, para Snyder (2016b), es importante distinguir las lenguas que contienen compuestos N+N en su léxico, pero que no permiten la creación de nuevos compuestos, de las lenguas que sí la permiten. En algunas lenguas, como el inglés, la composición es una estrategia muy productiva, mientras que, en lenguas como el francés o el español, se privilegian otras estrategias, como la derivación o la expresión mediante un sintagma nominal complementado por un sintagma preposicional. De este modo, en inglés, encontramos compuestos como (3.31) o (3.32),

¹⁰² Plag (2006) explica que las llamadas *Nuclear stress rule* y la *Compound stress rule* permiten capturar la diferencia de acentuación entre los sintagmas que se suelen acentuar al final y los compuestos que se suelen acentuar en el primer elemento (p. ej. el compuesto “blackboard” versus el sintagma “black board”). Referimos al lector al trabajo de Plag que trata no solo de las reglas sino también de las excepciones a estas reglas.

¹⁰³ Esta noción se exploró también en los trabajos Hualde (2007) y Rao (2015).

mientras que sus equivalentes en francés y en español suelen ser, o bien palabras derivadas, como en el caso de (3.31), o sintagmas nominales complementados por un sintagma preposicional, como en (3.32).

(3.31) *Apple tree* (en inglés)

Manzana árbol

“manzano” en español

“pommier” en francés

(3.32) *Sail boat* (en inglés)

Vela barco

“velero” o “barco de vela” en español

“voilier” o “bateau à voile” en francés

Esta ausencia o menor productividad de la composición en francés y en español no impide que encontremos compuestos de uso común, como (3.33), (3.34) en esas lenguas, pero según el análisis de Snyder (2001, 2012, 2016b, 2016a), esto no es suficiente para afirmar que la composición sea creativa en estas lenguas.¹⁰⁴

(3.33) *homme-grenouille*

hombre rana

“hombre provisto del equipo necesario para efectuar trabajos submarinos” (RAE, 2014)

(3.34) *mot-clé*

palabra clave

“palabra significativa o informativa sobre el contenido de un documento...” (RAE, 2014)

De hecho, como veremos en más detalle en la sección 3.3, Snyder (2001, 2012, 2016b, 2016a) es categórico en cuanto a la clasificación de las lenguas que permiten de manera creativa la composición. Sin embargo, encontramos autores como Arnaud (2003, 2016, 2018) que consideran que la composición en francés sí es productiva, como se desprende de los ejemplos en (3.35) y (3.36).

(3.35) *journal poubelle*

¹⁰⁴ En su trabajo, Snyder (2016a) insiste en el carácter binario de la creatividad de los procesos de formación de palabras: “A process of word formation either is, or is not, creative. Here ‘creative’ is being used in the sense of ‘the creative aspect of human language’: the fact that we can freely create a new sentence, potentially one that has never been used before, and reasonably expect a listener to understand it. A specific type of compound word formation is likewise creative, in a given language, if it is available for automatic, impromptu use whenever a new word is needed to fit the occasion” (p. 91)

- periódico basura
 “periódico de pobre calidad”
 (3.36) *camion poubelle*
 camión basura
 “camión que sirve para transportar la basura”

Otro argumento de Snyder (2016a) radica en la interpretación de los compuestos. Si una lengua es creativa, permite una variedad de significados para un mismo compuesto y, en el caso contrario, permite una sola interpretación clara (p. 91). Por ejemplo, siguiendo a Bauer (1978, pp. 83-84), Snyder afirma que el ejemplo (3.33) es un calco del inglés “*frogman*” y que, por esa razón, tiene el significado original de su equivalente en inglés. En cambio, en inglés, según el contexto, el mismo compuesto se podría interpretar de distintas maneras: un hombre que vende ranas, un hombre aficionada de las ranas, etc. Los ejemplos (3.35) y (3.36) son contraejemplos claros de esa afirmación, ya que su N2 se puede emplear con diferentes significados en otros compuestos del francés. Por lo tanto, podríamos acuñar un compuesto como (3.37) y tener diferentes interpretaciones posibles,¹⁰⁵ del mismo modo que se pueden interpretar ciertos compuestos de lenguas consideradas como creativas por Snyder. De manera similar, Noally (1990) subraya la ambigüedad de la estructura que reproducimos en (3.38).

- (3.37) *vélo poubelle*
 bicicleta basura
 “¿bicicleta que sirve para transportar la basura?” o “¿bicicleta de pobre calidad?”
 (3.38) *conférence sandwich* (Noally, 1990, p. 33)
 conferencia sándwich
 “¿conferencia breve entre dos otras conferencias?” o “¿conferencia sobre la hora de la comida?”

Considerando estos ejemplos, la noción de interpretación no parece tan categórica para determinar si una lengua permite o no la composición de manera creativa. Por último, otro elemento relacionado con la creatividad de los compuestos en las lenguas, según Snyder (2016a) y Namiki (1994), es la recursividad. Según la generalización de Namiki (1994), los N+N endocéntricos son creativos en una lengua solo si el proceso de creación es recursivo, como se ilustra en el ejemplo (3.39), adaptado de Snyder (2016a).

- (3.39) a. *flower book*
 flor libro

¹⁰⁵ Es posible que, a este respecto, el francés sea más permisivo que el español.

- “libro de flores”
- b. *flower* *book* *collection*
 flor libro colección
 “colección de libros de flores”

El inglés permite que un compuesto sirva de componente dentro de otro compuesto de manera recursiva. Si bien esto puede ocurrir también en ciertos compuestos deverbales del español, como (3.40), el fenómeno está más constreñido que en inglés y no se puede aplicar a los compuestos N+N.

- (3.40) a. parabrisas
 b. limpiaparabrisas

Volviendo a la noción de productividad en general, dentro de una misma lengua, ciertos tipos de compuestos pueden ser más productivos que otros. En inglés, por ejemplo, los compuestos N+N son más productivos que los compuestos deverbales. Por el contrario, en español y en las lenguas romances de manera general, los deverbales son más productivos que los N+N (Kornfeld, 2009; Fradin, 2009).

Guevara y Scalise (2009) afirman, en contra de la clasificación binaria de Snyder, que la noción de productividad es relativa y que la posibilidad de acuñar nuevos compuestos es inevitablemente gradual (p. 123). A continuación, indagamos en los contrastes de productividad de los compuestos entre las lenguas de nuestro estudio.

3.2 La composición en español, francés y euskera

Hasta aquí hemos presentado varias formas de compuestos de las lenguas del estudio para ilustrar algunas nociones preliminares, pero en la presente sección se describirán de manera más sistemática los compuestos N+N endocéntricos y deverbales y se aludirá también a otros tipos de compuestos cuando sea pertinente para compararlos con los N+N y los deverbales. Para estos dos tipos de compuestos, se presentarán, por una parte, sus características morfológicas y ortográficas y, por otra, su productividad y las relaciones semánticas y gramaticales entre sus componentes. Presentamos las tres lenguas del estudio de manera separada, empezando por el español.

3.2.1 Español

3.2.1.1 Compuestos N+N

En español, los compuestos endocéntricos subordinativos de tipo N+N se forman con el núcleo a la izquierda (p. ej., “**barco** pirata”). Aunque existen excepciones en cuanto a la direccionalidad (p. ej., “videojuego” en español, donde el núcleo “juego” aparece a la derecha), el orden antes mencionado es el más productivo (Moyna, 2006, p. 619). Según su grado de lexicalización, pueden escribirse como una sola palabra (p. ej., “bocacalle”, “telaraña”) o con los dos componentes separados por un espacio.¹⁰⁶ En cuanto al lugar de la flexión (el llamado *locus inflectionis*), cuando los dos elementos aparecen fusionados, la marca del plural se coloca al final de la palabra (p. ej., “telarañas”). Sin embargo, cuando los dos elementos aparecen separados, el español sigue la tendencia general que consiste en flexionar solamente el núcleo de los compuestos (Zwicky, 1985). Ciertos autores afirman que existe también la posibilidad de flexionar el elemento no-nuclear, aunque se trate de un uso marginal (Guevara & Scalise, 2009, p. 124), mientras que otros califican esta posibilidad de agramatical (Liceras, 2009, p. 26). Liceras añade que el núcleo es también el lugar indicado para colocar los afijos derivativos aumentativos o diminutivos, una afirmación validada por un grupo de hablantes nativos de español en un estudio experimental (Liceras & Klassen, 2019).

- (3.41) a. hombre rana (Guevara & Scalise, 2009, p. 124)
b. hombres rana
c. hombres ranas
- (3.42) a. mujer araña / hombre araña (Liceras, 2009, p. 26)
b. dos mujeres araña
c. *dos mujeres arañas
d. *hombre araño
e. perro policía
f. perrito policía / perritos policía / perrazos policía

Los ejemplos presentados en (3.41) y (3.42) ilustran la posición respectiva de los autores en cuanto a la posibilidad de flexionar el elemento no-nuclear en número, mientras que los ejemplos (3.42) muestran la posibilidad de combinar los morfemas flexivos y derivativos, pero solo en el

¹⁰⁶ El guion se usa en los compuestos coordinativos (p. ej., “director-presentador”, “lectura-escritura”) que, como hemos mencionado anteriormente, quedan fuera del marco del estudio.

elemento nuclear. Por su parte, la NGLE añade que ciertos compuestos parecen permitir de manera más frecuente la flexión de número en sus dos elementos, como ilustran los ejemplos de (3.43).

- (3.43) a. asunto clave (RAE & ASALE, 2010a, p. 193).
b. asuntos clave ~ asuntos claves
c. copia pirata
d. copias pirata ~ copias piratas

Según la NGLE, la presencia de la flexión en el segundo elemento depende de si se interpreta la expresión como un compuesto o como un grupo sintáctico. Dado que la mayoría de los compuestos N+N en español exhiben una relación similar a la que existe entre el nombre y el adjetivo restrictivo en un sintagma nominal, resulta pertinente comparar la flexión en los compuestos de tipo N+A con la de los de tipo N+N. Los ejemplos en (3.44) ilustran cómo, en estos compuestos N+A endocéntricos, la flexión se realiza en los dos elementos, como ocurriría en un sintagma nominal formado por un nombre y un adjetivo y la ausencia de concordancia daría lugar a un compuesto agramatical.¹⁰⁷

- (3.44) a. una mesa redonda
b. dos mesas redondas
c. *dos mesas redonda / *dos mesas redondo
d. un bulto redondo
e. dos bultos redondos
f. *dos bultos redondo

En cambio, la situación se complica cuando observamos la flexión en los compuestos N+A exocéntricos. Los ejemplos en (3.45) ilustran que ciertos compuestos N+A permiten la concordancia de número con los dos elementos del compuesto, pero que otros como en (3.46) no la permiten. Es más, en (3.46), la concordancia en género entre el determinante y el sustantivo no se realiza, mientras que el sustantivo y el adjetivo concuerdan entre sí (RAE & ASALE, 2010a, p. 199).¹⁰⁸

- (3.45) a. un cabeza rapada

¹⁰⁷ El “bulto redondo” es un compuesto del campo de la escultura que se refiere a una “obra escultórica exenta, y por tanto visible por todo su contorno.” (RAE, 2014)

¹⁰⁸ El ejemplo (3.46) no es agramatical cuando consideramos relación pública como un simple sustantivo modificado por un adjetivo, pero se puede usar con el significado del compuesto exocéntrico que hace referencia a una “persona que desempeña las relaciones públicas” (DLE).

- b. una cabeza rapada (referente femenino-singular)
 - c. unos cabezas rapadas
 - c. unas cabezas rapadas (referente femenino-plural)
 - d. *unos cabezas rapada
- (3.46)
- a. *un relación pública
 - b. *una relación pública
 - c. un relaciones públicas
 - d. una relaciones públicas (referente femenino-singular)
 - e. unas relaciones públicas (referente femenino-plural)
 - f. unos relaciones públicas

Esta constatación sobre la flexión en los compuestos [N+A] nos permite constatar que la frontera borrosa que existe entre algunos compuestos [N+A] y [N+N] podría ser una fuente de confusión en cuanto a la flexión presente en el segundo elemento de los N+N. Esto nos lleva a profundizar en la cuestión de las relaciones semánticas posibles entre los elementos de los compuestos en español.

La mayoría de los compuestos de tipo N+N en español son atributivos, es decir que el N2 de esos compuestos cumple la función de atributo. Sin embargo, existen también N+N no atributivos (o predicativos, como los llama de Bustos Gisbert, 1986), como “coche cama”, donde “cama” no designa un atributo del coche, sino una relación de posesión. De Bustos Gisbert (1986) también menciona la relación instrumental con ejemplos como “cuerda-carro” (N1 que sirve para N2), la relación causa-efecto con ejemplos de plantas curativas como “yerbarreúma” (N1 que produce N2 o que evita N2) y, finalmente, una relación que incluye todo lo demás con ejemplos como “pájaro moscón” (N1 se alimenta de N2) (p. 226). Rainer y Varela (1992, p. 126) mencionan otros ejemplos como “fútbol sala”, “coche bomba” y “año luz”, pero afirman que las relaciones semánticas expresadas en esos compuestos son poco productivas y se deben principalmente a préstamos de otras lenguas. Encontramos otras relaciones, sobre todo en compuestos altamente lexicalizados como “telaraña” y “bocacalle”, que han sido poco estudiadas en español, ya que se consideran poco productivas. Una excepción a esta tendencia es el trabajo de Toquero-Pérez (2018), en el cual se estudian las relaciones semánticas de los N+N subordinativos del español, siguiendo la clasificación de Jackendoff (2011, 2016), inicialmente desarrollada para los

compuestos del inglés.¹⁰⁹ Reproducimos en la Tabla 3.1 las categorías que Toquero-Pérez encontró en su corpus de compuestos N+N del español.

¹⁰⁹ La clasificación de Jackendoff se basa en la Arquitectura Paralela. Para más detalles sobre esta propuesta teórica, referimos al lector a los trabajos de Jackendoff (2011, 2016; entre otros). Para este trabajo, nos limitamos a presentar las relaciones semánticas expuestas en este marco teórico para la presente discusión. Mantenemos el nombre de las funciones en inglés para facilitar al lector su identificación en las obras de referencia correspondientes.

Tabla 3.1. Las funciones básicas de Jackendoff (2011, 2016) utilizadas en los compuestos N+N subordinativos del español a partir de los datos de Toquero-Pérez (2018, pp. 37-38)

Relación semántica	Ejemplos en español	<i>n</i>
<i>SIMILAR</i> N1 es similar a N2	cabeza buque, coche oruga, tren bala	76
<i>PF</i> (función propia) N1 cuya PF es la de N2	hombre orquesta, lápiz láser, mecánico dentista	38
<i>X serves as Y</i> N1 sirve de N2	canción protesta, perro guardián, sombrero paraguas	33
<i>BE (AT/IN/ON)</i> N1 que está ubicado en/sobre N2; N1 que tiene N2 sobre él; N1 que tiene lugar en el tiempo N2; N1 cuya PF es estar en N2; N1 cuya PF es tener N2 sobre él	balón reloj, gas ciudad, zona euro	28
<i>HAVE</i> N1 que tiene N2; N1 que N2 tiene	bandera pirata, carril bus, coche patrulla	25
<i>COMP</i> N1 tiene N2 en su composición	avión cohete, barco tanque, flota pirata	15
<i>PART</i> N1 es parte de N1	aguanieve, mochila arnés, papel carbón	15
<i>CAUSE</i> N1 está causado por N2	efecto placebo, guerra narco, virus zombie	14
<i>BE</i> N1 es un N2	abeja reina, perro vampiro, príncipe rana	7
<i>ARGUMENT SCHEMA</i> Un N1 de/por N2	abordaje pirata, ataque pirata, bote salvavidas	5
<i>MAKE</i> N1 hecho por N2; N1 hace N2	fiesta pirata, horno microondas, telaraña	5
Total		261

De las funciones que Jackendoff enumera para los compuestos del inglés, cuatro no aparecen en la tabla, ya que Toquero-Pérez no encontró ningún compuesto en español para las funciones *CLASSIFY*, *KIND*, *MADE FROM* y *PROTECT*. La ausencia de compuestos que corresponden a la función *CLASSIFY* se debe sin duda al tamaño reducido del corpus, ya que los ejemplos proporcionados por Jackendoff en inglés (*beta cell*, *X-Ray*, *Molotov cocktail*) tienen sus equivalentes en español (célula beta, rayos X, cóctel molotov). Con respecto a las funciones *MADE FROM* y *PROTECT*, según los datos de Toquero-Pérez, el español expresa estas relaciones exclusivamente con la construcción N+de+N, como en “aceite de girasol”, “estofado de carne” o “caña de azúcar”, en el caso de la función *MADE FROM*, y “botas de lluvia/agua”, cinturón de castidad” y “gafas de sol” para la función *PROTECT*. En cuanto a la función *KIND* que aparece en ejemplos del inglés como “*pine tree*”, “*gemstone*”, “*limestone*” (“pino”, “gema”, “caliza”), Toquero-Pérez no encontró ejemplos ni de tipo N+N, ni de tipo N+de+N. Volviendo a los datos que aparecen en la Tabla 3.1, la variedad de las relaciones muestra que los compuestos N+N endocéntricos del español no son meramente atributivos (lo que correspondería, a grandes rasgos, a las funciones *SIMILAR*, *PF*, *BE*), sino que pueden expresar casi todas las mismas funciones básicas que Jackendoff encontró con los compuestos del inglés. Ahora bien, la última columna de la Tabla 3.1 muestra que no todas las funciones aparecen con la misma frecuencia, lo que indica que unas son más productivas que otras en español. De manera general, se considera que las relaciones semánticas de los compuestos en español están más restringidas que en las lenguas germánicas (Rainer, 2016). Rainer afirma también que el francés exhibe menos restricciones semánticas que el español en el proceso de formación de los N+N (p. 2621).

En cualquier caso, en español, el proceso de formación de los compuestos N+N no atributivos no es tan productivo como la derivación (“manzano” y no “árbol manzana”) o el uso de estructuras con preposición [N+prep+N] (“barco de vela” y no “barco vela”).

Volviendo a la relación semántica de atribución, se observa que, en español, ciertos sustantivos (“clave”, “cumbre”, “pirata”, “estrella”, “fantasma” y “relámpago”, entre otros) tienen un uso muy productivo como segundo elemento de compuestos N+N, ya que estos denotan una característica prototípica bien asentada (Serrano Dolader, 2018). Por ejemplo, “clave” denota importancia y se puede yuxtaponer a una multitud de sustantivos de manera productiva con este significado: “capítulo clave”, “palabra clave”, “asignatura clave”, etc. Rainer y Varela hablan de

un proceso de evolución hacia la función de adjetivo de algunas de las voces que acabamos de mencionar (Rainer & Varela, 1992, p. 126).¹¹⁰

A modo de ejemplo, exploramos más en detalle el proceso de evolución de la voz “pirata”, un elemento que aparece como N2 en varios ejemplos de la Tabla 3.1 y que forma parte del compuesto “barco pirata”, que aparece a menudo en los trabajos de adquisición sobre los compuestos N+N (Liceras, 2003, 2009; Liceras et al., 2002; Liceras & Díaz, 2000b, 2000a, 2000b; Pomerleau, 2000; Trías & Villanueva, 2011). De las tres primeras acepciones de esta voz recopiladas en la presente edición del *DLE* (RAE, 2014) y que reproducimos en (3.47), es llamativo el hecho de que las dos primeras acepciones correspondan a la categoría de adjetivo.

- (3.47) 1. adj. Perteneciente o relativo al pirata o a la piratería.
2. adj. Ilegal, que carece de la debida licencia o que está falsificado.
3. m. f. Persona que, junto con otras de igual condición, se dedica al abordaje de barcos en el mar para robar. (RAE, 2014)

Una consulta del *Nuevo tesoro lexicográfico de la lengua española* (NTTLE; RAE, 2001b) nos permite constatar que, en la evolución de esa voz, la primera acepción adjetival solo aparece en el *Diccionario de la lengua española 15ª ed.* (RAE, 1925). En el *Diccionario de la lengua castellana por la Real Academia Española 14ª ed.* (RAE, 1914, p. 805), las dos únicas acepciones, que reproducimos en (3.48), corresponden a la categoría de sustantivo:

- (3.48) 1. m. Ladrón que anda robando por el mar.
2. fig. sujeto cruel y despiadado que no se compadece de los trabajos de otro.

La segunda acepción –“figurativa” y que las ediciones anteriores llamaban “metafórica”– parece ser la responsable de la posterior evolución hacia la función de adjetivo. A su vez, la voz “cumbre” parece haber seguido un proceso similar. En (3.49), enumeramos las diferentes etapas de evolución de esta voz.

¹¹⁰ La adjetivación de un sustantivo es un proceso de recategorización en el cual este adquiere las funciones sintácticas de un adjetivo, como en “muy señora” o “muy hombre”, conservando solo “alguna propiedad culturalmente relevante que caracteriza a ciertos grupos de personas” (RAE & ASALE, 2011, p. 248). La adjetivación no se hace exclusivamente tomando como punto de partida la clase de los sustantivos, véase, por ejemplo, el trabajo de Pato (2010) sobre la recategorización del adverbio “medio” en la clase de los adjetivos.

- (3.49) a. sustantivo que, además de su acepción principal (“cima o parte superior del monte”), cuenta con una acepción figurativa (“la mayor elevación de una cosa o último grado a que puede llegar”) hasta la 21ª edición del diccionario (RAE, 1992);
- b. aparece una segunda acepción figurativa (“reunión de máximos dignatarios nacionales o internacionales para tratar asuntos de especial importancia”) en la 21ª edición del diccionario (RAE, 1992);
- c. aparece la sub-entrada “conferencia cumbre” en la edición siguiente (RAE, 2001a) con la acepción de “[Conferencia] celebrada entre jefes de Estado o de Gobierno para consultar o decidir cuestiones importantes”;
- d. en la última edición (RAE, 2014), aparece una acepción propiamente adjetival: “adj. Dicho de una cosa: Que tiene la máxima perfección o importancia en su género”.

Tomando estas dos palabras como ejemplos, podemos constatar que para que una palabra se pueda combinar con otro sustantivo formando un compuesto N+N en español, esta tiene que haber iniciado un cierto camino hacia la recategorización como adjetivo.¹¹¹ Este patrón, por el que el N2 aporta un valor metafórico al N1, parece ser el más productivo para los N+N del español. Se trata de una estructura compositiva parecida a la de los compuestos N+A en el sentido de que el N2 modifica al N1 de manera comparable a la del adjetivo de los compuestos N+A. Cabe mencionar que autores como García-Page (2011) prefieren denominar “pseudocompuestos” a las secuencias como “momento cumbre”, en la que el N2 cumple una función más adjetival. En el marco de nuestro trabajo, si bien entendemos la necesidad de subrayar esta diferencia desde un punto de vista descriptivo, creemos necesario incluir todas las construcciones que correspondan formalmente a compuestos N+N, ya que, desde la percepción del hablante (no experto), esas diferencias pueden no ser tan salientes.

En resumen, de manera general, los compuestos N+N endocéntricos subordinativos del español se forman con el núcleo a la izquierda, se flexionan solamente en su núcleo y son poco productivos. Por último, el tipo de N+N más productivo es aquel en el que el N2 tiene un valor cercano al de un adjetivo, punto que profundizaremos en el análisis de los compuestos N+N en francés.

¹¹¹ Volveremos con más detalles sobre este proceso al describir los compuestos N+N en francés (sección 3.2.2.1), ya que se trata de un tema bien desarrollado por Arnaud (2018) en francés, pero solo superficialmente para el español (Rainer & Varela, 1992; Serrano Dolader, 2018), con la excepción de García-Page (2011).

3.2.1.2 Compuestos deverbales

El mecanismo para formar los compuestos deverbales es muy productivo en español (J. W. Harris, 1987; Lang, 1991; Lardiere & Schwartz, 1997). Los compuestos deverbales, como por ejemplo “abrelatas”, son generalmente considerados exocéntricos, de tipo verbo-objeto y se les atribuye el género masculino por defecto.

En cuanto a la morfología y la ortografía, se forman a partir de un verbo, que adopta la forma de la tercera persona del singular del presente de indicativo, seguido de su objeto. El verbo y el objeto aparecen formando una sola palabra, sin espacio ortográfico ni guion que separe ambos elementos. El objeto aparece generalmente en plural y recibe, de este modo, un valor genérico.

Basándose en los compuestos deverbales registrados en la 20ª edición del *Diccionario de la lengua española* (RAE, 1984), Pérez Lagos (1986, p. 27) afirma que, en la mayoría de los casos, el verbo empleado pertenece a la primera conjugación (-AR). En un estudio actualizado a partir de la última edición del *Diccionario de la lengua española* (DLE; RAE, 2014), Lacroix (en preparación) confirma que esta tendencia sigue vigente, ya que más del 90 % de los compuestos deverbales registrados en el DLE se forman a partir de la primera conjugación, mientras que solo el 6 % se forman a partir de verbos de la segunda (-ER) y apenas un 3 % a partir de la tercera (-IR).¹¹² Al contrario de lo que veremos con el francés, la forma del primer elemento no presenta excepciones y siempre corresponde a la grafía de la forma de tercera persona del singular.

Con respecto a la morfología del segundo elemento, si bien el requisito de la -s genérica parece ser la opción por defecto, esta regla no está exenta de excepciones. Val Álvaro (1999, p. 4798) menciona tres tipos de excepciones. En primer lugar, la -s genérica puede denotar un cambio de significado como en los ejemplos (3.50) y (3.51):¹¹³

¹¹² Los 15 verbos más frecuentes en Pérez Lagos (1986, p. 27) eran: guardar, matar, portar, sacar, saltar, tirar, cortar, pasar, tapar, romper, tragar, picar, quitar, parar y tornar. En el registro actualizado de Lacroix (en preparación), a pesar de las 204 nuevas incorporaciones al DRAE, esta lista se mantiene casi inalterada, salvo por “lavar” que entra a formar parte de esta lista (posición 11), mientras que “tornar” queda relegado a la posición 23 y “parar” llega *ex aequo* con “chupar” en la posición 15.

¹¹³ Estos cambios semánticos en función de la presencia o ausencia de la -s final parecen variar de una zona dialectal a otra. Buenafuentes de la Mata (2017) observa al respecto que “esta posibilidad de distinguir significados a partir de la adición de la -s se da en muchos más casos [en el español de América].” (pp. 682-683)

- (3.50) catavinos (persona) / catavino (instrumento) (Val Álvaro, 1999, p. 4798)
(3.51) sacabalas (baqueta) / sacabala (pinza del cirujano)

Además, cuando el complemento es una entidad única, como en (3.52), o una entidad no contable o colectiva, como en (3.53), se tiende a preferir la forma sin -s générica.¹¹⁴

- (3.52) girasol, portabandera (Val Álvaro, 1999, p. 4798)
(3.53) guardarropa, cortacésped

En cuanto al valor semántico del referente de los deverbales, según la NGLE (RAE & ASALE, 2009, p. 201), el más productivo es el de instrumento, como en (3.54), seguido del de agente (3.55). También se pueden usar deverbales para designar acciones o procesos, como se muestra en (3.56).

- (3.54) cortaúñas
(3.55) limpiabotas
(3.56) besamanos

Como explica Moyna (2011), todas estas relaciones semánticas aparecen registradas ya en el español medieval, y, desde las etapas iniciales de la lengua, se registran deverbales con referentes tanto animados (~agentes) como no animados (~instrumentales, ~eventos), si bien la frecuencia relativa de cada uno de estos rasgos varía a lo largo de la historia. Según un análisis de los 75 deverbales incorporados a la 23ª edición del DLE (y sus posteriores modificaciones hasta 2019), el 62,7 % de las nuevas incorporaciones tiene un valor instrumental (instrumento, planta, lugar), mientras que el 29,3 % hacen referencia a un agente (humano o animal), y el 5,3 % alude a un evento (Lacroix, en preparación). Dos de esas nuevas incorporaciones se usan tanto para referirse

¹¹⁴ Aparte de esos tres tipos de excepciones mencionados por Val Álvaro, existen otras excepciones aisladas y que se han asentado en el uso con o sin -s. Por ejemplo, encontramos un contraste entre el bien asentado “cumpleaños” que lleva la -s générica y una de las adiciones de la edición 2019 del DLE “cumplemés” que no la lleva (RAE, 2019). Esta oposición no se puede explicar ni por contraste con la variante “cumplemeses” (también usada con el mismo sentido), ni con el hecho de que mes sea una entidad única y tampoco colectiva o no contable. De manera similar, otra novedad de la actualización del 2019, la adición de la entrada besapié (sin -s), como sinónimo de besapiés resulta difícil de categorizar en las excepciones de Val Álvaro.

a un instrumento como a un agente (limpiavidrios, limpiacristales).¹¹⁵ Si bien los deverbales documentados en el diccionario no son el reflejo exacto de su uso en la actualidad, al observar las incorporaciones recientes, podemos constatar los patrones más productivos. En este caso, el análisis de las nuevas entradas corrobora la tendencia documentada en varios tratados sobre los compuestos: los deverbales se usan principalmente con un valor instrumental, en menor medida con un valor agentivo y de manera marginal para hacer referencia a un evento.

Con respecto a las relaciones gramaticales entre el verbo y el sustantivo, la más productiva es la de paciente (objeto directo), en menor medida la de lugar (complemento circunstancial de lugar) y de manera marginal la de agente (sujeto), como muestran los ejemplos (3.57), (3.58) y (3.59), respectivamente.¹¹⁶

(3.57) lavaplatos, abrelatas, espantapájaros

(3.58) andarríos, corre caminos, trotacalles

(3.59) rodapié, vuelapié, vuelapluma

Se considera generalmente que el caso prototípico de los deverbales en español es el de paciente, que cumple la función de objeto directo del verbo. De nuevo, el análisis de las nuevas incorporaciones en la 23ª edición del DLE nos confirma esta tendencia: en el 94,7 % de los casos, el sustantivo actúa como objeto directo y, en el resto de los casos, como un complemento circunstancial de lugar (Lacroix, en preparación).¹¹⁷

Para concluir el análisis sobre los deverbales en español, cabe mencionar que estos compuestos compiten con otros procesos de creación léxica, particularmente la derivación mediante el sufijo -dor. Efectivamente, existen varios pares donde el compuesto verbal tiene un significado similar al derivado en -dor o en -ero, por ejemplo: “abrebotellas ~ abridor”,

¹¹⁵ Como algunas de las nuevas incorporaciones al diccionario, estas entradas representan usos propios de ciertas áreas geográficas. En este caso, los dos deverbales son equivalentes, salvo que el uso de “limpiacristales” está documentado en España, y “limpiavidrios” en Argentina, Chile, Colombia, Costa Rica, México, Nicaragua y Paraguay.

¹¹⁶ Los ejemplos proceden de tratados sobre los compuestos (de Bustos Gisbert, 1986; Rainer & Varela, 1992) o directamente de la NGLE (RAE & ASALE, 2009, p. 201). Hay que subrayar que los ejemplos con complementos de lugar o de agente son más difíciles de encontrar y parecen poco productivos en la actualidad, por lo que pueden resultar difíciles de interpretar incluso para un hablante nativo.

¹¹⁷ Los cuatro deverbales con complemento circunstancial de lugar son: corre pasillos, corre calle, vuelacerca y pasaboca(s).

“afilalápices ~ afilador”, “lavavajillas ~ lavadora”, “sacacorchos ~ descorchador”, “portatarjetas ~tarjetero”(Serrano Dolader, 2018, p. 316). Esta competencia entre las dos formas puede afectar a la productividad de estos deverbales frente a otros que no tienen equivalentes derivados con un significado tan cercano (por ejemplo, el término “guardador” no es aplicable para hacer referencia a un “guardarropa”). A pesar de ello, de manera general, los compuestos deverbales siguen siendo un tipo de compuesto muy productivo en español.

En resumen, los compuestos deverbales del español se forman con el verbo al inicio, seguido por un sustantivo. El sustantivo que cumple generalmente la función de objeto directo del verbo y lleva una -s final que le da un valor genérico. El valor semántico más frecuente en este tipo de compuesto es el de instrumento, también se usa a menudo para hacer referencia a un agente animado.

3.2.2 Francés

Los compuestos nominales en francés siguen en gran medida los patrones de formación de sus equivalentes en español. Por lo tanto, presentaremos de manera más sintética los aspectos similares entre el francés y el español y nos detendremos más en las divergencias entre las dos lenguas.

3.2.2.1 Compuestos N+N

Como en español, los compuestos N+N endocéntricos subordinativos se forman con el núcleo a la izquierda (p. ej., “*bateau pirate*”, “barco pirata”) (Arnaud, 2016, p. 73). Aparte de los llamados compuestos neoclásicos (los llamados “*composés savants*” en francés), cuyo núcleo aparece a la derecha (p. ej., “*ludothèque*”, “ludoteca”), las excepciones en cuanto a la direccionalidad se encuentran principalmente en neologismos tomados de otras lenguas (p. ej., “*webmestre*”, del inglés “*webmaster*”, es decir, “administrador de sitio web”). Las reglas que guían la ortografía de los compuestos en francés son, de manera general, menos nítidas que en español. La reforma ortográfica de 1990 recomienda una mayor estandarización de las reglas de uso (Conseil supérieur de la langue française, 1990), pero las excepciones y algunas tendencias en el uso hacen poco previsible la ortografía de un compuesto dado. Picone (1996, p. 112) considera que, dada la gran variación en el uso y el hecho de que el plural no es perceptible en la lengua oral en francés, las reglas establecidas por los gramáticos y lingüistas no reflejan necesariamente las

intuiciones de los hablantes. Con esto en mente, a continuación, exploramos estas reglas ortográficas sobre el uso del guion y del plural en los compuestos.

En cuanto al uso del espacio y del guion, si bien se recomienda la fusión sin espacio de los componentes en varios tipos de compuestos (los formados por “*contre-*” o prefijos latinos como “*extra-*”, “*intra-*”, etc.) y la unión mediante un guion para los compuestos N+N coordinativos, las pautas para los compuestos N+N endocéntricos no lexicalizados son menos claras y, a menudo, las obras de referencia recopilan formas con o sin guion, siguiendo criterios, a veces, difícilmente identificables. Por ejemplo, la gramática Grevisse menciona que el uso del guion está bien asentado en el compuesto “*bébé-éprouvette*” (Grevisse & Goosse, 2008, p. 194), pero en el diccionario *Le Grand Robert* (2017), la grafía con el espacio también aparece como posible. Este diccionario hace una distinción en cuanto al uso del guion entre lo que califica de “sustantivos en aposición” y los compuestos. Así, para el sustantivo “radar”, *Le Grand Robert* (2017) presenta una lista de sustantivos en aposición (3.60) y una serie de compuestos (3.61). El matiz entre las dos listas resulta difícil de percibir, por lo que profundizaremos el tema al hablar de las relaciones semánticas entre los componentes tras los detalles morfológicos.

- (3.60) *système radar, écran radar, sondeur radar, balise radar*
sistema radar, pantalla radar, sonda radar, baliza radar
- (3.61) *contrôle-radar, émetteur-radar, réflecteur-radar*
control-radar, emisor-radar, reflector-radar

En cuanto a la flexión de número, Villoing (2012) subraya que hay poca regularidad y que, en algunos casos, un compuesto puede aparecer flexionado de dos maneras distintas por diferentes autores (p. 53). Sin embargo, Villoing afirma que el criterio principal que guía la flexión de los compuestos N+N es la relación semántica entre sus elementos: en los coordinativos (3.62) y los atributivos (3.63) se suelen flexionar los dos elementos, mientras que en los subordinativos (3.64) se flexiona solo el núcleo, es decir, el primer elemento (3.64).

- (3.62) *des boulangers-pâtisseries, des moissonneuses-batteuses, des bars-tabacs*
unos panaderos-pasteleros, unas segadoras-trilladoras, unos bares-tabacos
“unos panaderos-pasteleros, unas segadoras-trilladoras, unos bares-tabaquerías”
- (3.63) *des poissons-chats, des hommes-grenouilles, des oiseaux-mouches, des choux-fleurs*
unos peces-gatos, unos hombres-ranas, unos pájaros-moscas, unos coles-flores
“unos peces gato, unos hombres rana, unos pájaros mosca, unas coliflores”

(3.64) *des timbres-poste, des pauses-café, des wagons-poste, des bébés-éprouvette.*
unos sellos-correo, unas pausas-café, unos vagones-correo, unos bebés-probeta
“unos sellos (postales), unas pausas de café, unos vagones postales, unos bebés probeta”

Otras fuentes (*Antidote 9*, 2018) mencionan que, si el segundo elemento es un complemento del primero, como en “timbre-poste”, solo el primero concuerda en número (“un timbre-poste”, “des timbres-poste”), lo que coincide con la descripción de Villoing. En cuanto a los ejemplos similares a los que aparecen en (3.63) y algunos de (3.62) y que esta fuente denomina bajo la expresión *apposition attachée* (aposición adjunta, atada), la concordancia en número puede variar según criterios que pueden parecer arbitrarios, como si el compuesto está formado por dos oficios.¹¹⁸

Como podemos ver, las gramáticas y diccionarios del francés emplean una variedad de términos para clasificar diferentes construcciones que en nuestro trabajo consideramos como compuestos. Arnaud (2003) comenta que, en ocasiones, se denomina a los compuestos N+N como compuestos apositivos o compuestos por aposición, dos expresiones que el autor no considera apropiadas porque aluden a un proceso puramente sintáctico y distinto de la composición. Arnaud prefiere hablar de yuxtaposición o, simplemente, de estructura NN (p. 3). Coincidimos con Arnaud en que el término de aposición no es el más adecuado, ya que se trata de una noción polisémica que los gramáticos del francés usan de manera variable (Grevisse & Goosse, 2008, p. 425). Por ejemplo, las llamadas *appositions détachés* (aposiciones sueltas, desatadas), como en (3.65), se alejan de lo que consideramos como un compuesto. Sin embargo, lo que las gramáticas denominan “aposiciones atadas” (*apposition attachée*), como los ejemplos de (3.66), sí que corresponde en gran medida con nuestra definición de compuesto. Entonces, siguiendo a Arnaud, evitamos el término “aposición” y consideramos que varias de las estructuras consideradas como *appositions attachées* son, en realidad, compuestos.

(3.65) *les albatros, vastes oiseaux des mers* (Grevisse & Goosse, 2008, p. 425)

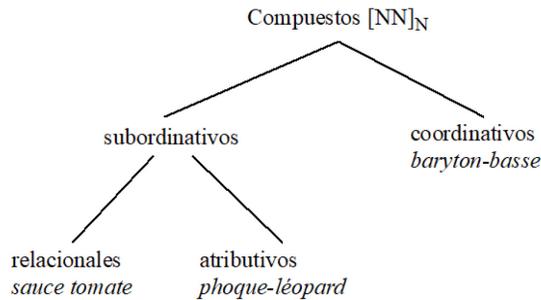
¹¹⁸ En realidad, el criterio aquí parece ser más bien la coordinación, ya que cuando se combinan dos oficios, se obtiene un compuesto coordinativo, como muestran los ejemplos citados en *Antidote* “des sapeurs-pompiers” y “des boulangers-pâtisseries”, que también enumera Villoing en su lista de coordinativos. Si bien *Antidote* formula algunas generalizaciones en la guía sobre los plurales de los compuestos, menciona que no hay reglas absolutas que dicten la concordancia de los sustantivos o adjetivos en los compuestos y que el uso no es siempre fijo o coherente, por lo que se aconseja mirar en un diccionario o usar una herramienta de corrección automática como *Antidote*.

	los albatros, vastas aves	de los mares	
(3.66) a.	<i>mots</i>	<i>clés</i>	(Grevisse & Goosse, 2008, p. 429)
	palabras	claves	
b.	<i>cas</i>	<i>limites</i>	
	casos	límites	
c.	<i>prix</i>	<i>choc</i>	
	precios	choque	
d.	<i>voyages</i>	<i>éclair</i>	
	viajes	relámpago	
e.	<i>hommes-</i>	<i>cheval</i>	
	hombres	caballo	

Volviendo a la flexión del plural, los ejemplos en (3.66) aparecen en plural para ilustrar la variación en la marca del plural en el segundo elemento. Según Grevisse y Goose (2008, p. 425) esta diferencia en la pluralización del segundo elemento es puramente gráfica, ya que en los casos donde el plural sería audible como en (3.66), el segundo elemento suele permanecer en singular (*cheval/*chevaux*). Este contraste nos lleva a hablar de una diferencia importante entre el francés y el español en cuanto a la ortografía de que afecta a los compuestos de manera particular. El sistema ortográfico del francés es considerado como opaco, mientras que el del español es considerado transparente (Lallier et al., 2014). Esto significa que, para la morfología flexiva de los compuestos, si bien la diferencia entre “hombres rana” y “hombres ranas” es perceptible en español, en francés no se percibe ninguna diferencia entre “*hommes grenouille*” y “*hommes grenouilles*”. Que sepamos, no hay estudios sobre el vínculo entre la opacidad ortográfica del francés y la flexión del plural en los compuestos, pero consideramos que una parte de la inestabilidad de esa flexión podría proceder de esta realidad.

Los compuestos N+N del francés muestran una inestabilidad evidente en cuanto a la ortografía y la morfología. Si bien es probable que esta inestabilidad esté relacionada a la opacidad de la ortografía, también se debe al hecho de que las reglas ortográficas están a menudo basadas en clasificaciones que varían de un autor a otro. En el marco de nuestro trabajo, retomamos la clasificación que hemos usado para el español, es decir, la clasificación de Arnaud (2016, p. 73), que reproducimos en (3.67) y que se basa en las relaciones entre los componentes de los compuestos.

(3.67)



En la rama “subordinativos”, para diferenciar los atributivos de los relacionales, Arnaud utiliza la prueba en (3.68), una prueba que no funciona con los compuestos relacionales.¹¹⁹

(3.68) Un N1N2 es un N1 es un (como) un N2

Un phoque-léopard est un phoque qui est comme un léopard
"una foca leopardo es una foca que es como un leopardo"

En contraste con el español que presenta principalmente compuestos N+N endocéntricos de tipo atributivo, el francés presenta una variedad de relaciones. A partir de un corpus de 809 compuestos en francés, Arnaud (2003) estableció una lista de 54 relaciones semánticas distintas entre los componentes de los compuestos N+N. A esta lista, a lo largo de los años, Arnaud añadió 140 compuestos y 4 nuevas relaciones, para llegar a una clasificación de 58 relaciones (Arnaud, 2016, p. 82). Arnaud afirma que su clasificación es compatible con la taxonomía de Jackendoff (2016) que introducimos para presentar los compuestos N+N del español (pp. 84-85). Reproducimos en (3.69) las cinco relaciones más frecuentes del corpus de Arnaud (indicamos el número de ocurrencias de cada relación entre paréntesis).¹²⁰

(3.69) a. el elemento no nuclear es el objetivo/propósito (general) del núcleo (284)

carte réponse
tarjeta respuesta
“tarjeta de respuesta”

¹¹⁹ Cabe mencionar que, en la lista de 58 relaciones de Arnaud, el ejemplo “*poisson-scie*” (pez sierra) que sirve para ilustrar la relación “el elemento no nuclear (concreto, discreto, metafórico) es una parte del núcleo” es interpretado por otros autores como un atributivo. Podemos considerar que el cuerpo entero del pez se parece a una sierra, por lo que se clasificaría como un compuesto atributivo, pero se puede considerar también que la sierra es una parte del pez y, en este caso, formaría parte de los compuestos subordinativos relacionales. Como se puede ver, esta clasificación resulta parcialmente subjetiva y la frontera entre ciertas relaciones atributivas y no atributivas es borrosa.

¹²⁰ La lista completa de las 58 relaciones semánticas aparece en el apéndice de Arnaud (2016).

- b. el elemento no nuclear es la posición del núcleo en el ambiente (espacio) (67)
enfant bulle
 niño burbuja
 “niño burbuja”
- c. el elemento no nuclear es el fenómeno/proceso/modo de acción/fuente de energía/ambiente que permite al núcleo funcionar o existir (59)
demi pression
 mitad presión
 “caña de cerveza”
- d. el elemento no nuclear (concreto o discreto) es una parte del núcleo (59)
tailleur pantalon
 traje pantalón
 “traje (de) pantalón”
- e. el elemento no nuclear es el origen (general) del núcleo (57)
arrêt maladie
 baja enfermedad
 “baja de enfermedad”

En el otro extremo de la lista, encontramos relaciones poco frecuentes que tienen una o dos ocurrencias, como las que reproducimos en (3.70):

- (3.70) a. el elemento no nuclear (abstracto, continuo) tiene el núcleo como soporte (2)
carte mémoire
 tarjeta memoria
 “tarjeta de memoria”
- b. el elemento no nuclear es donde fue identificado el núcleo (1)
facteur rhésus
 factor Rhesus
 “factor Rh”¹²¹

Si bien se puede observar que los compuestos N+N endocéntricos del francés presentan una variedad de relaciones, parece claro que algunas de ellas no son muy productivas y se pueden atribuir a distintos factores, como la influencia de otras lenguas (el inglés, de manera particular; Arnaud, 2003, p. 138). Sin embargo, el gran número de compuestos formados siguiendo ciertas relaciones, como las presentadas en (3.69), muestra la productividad de ciertos patrones de formación de los compuestos en francés, la cual parece levemente superior a los patrones no atributivos observados en español.

¹²¹ El macaco Rhesus es una especie de primate en el que fue identificado el llamado “factor Rh” del grupo sanguíneo.

En cuanto al proceso de adjetivación de ciertos N2 que abordamos en la sección sobre los compuestos N+N del español, se puede observar un fenómeno similar en francés. Arnaud (2018) estudió los sustantivos que aparecen en posición N2 de las estructuras N+N en francés para diferenciar a los sustantivos adjetivados de los sustantivos que funcionan como N2 de un compuesto. Su punto de partida es el sustantivo “*phare*” (faro) y el compuesto “*bateau phare*” (buque faro). Si bien “*phare*” tiene un valor composicional en “*bateau phare*”, Arnaud expone que su valor es más adjetival en expresiones como en las que reproducimos en (3.71). Si bien estas expresiones no se pueden traducir literalmente al español, resulta interesante destacar que, en todos los casos, “*phare*” se traduciría por el sustantivo “estrella”, un sustantivo que también muestra una tendencia a la adjetivación en español (Rainer & Varela, 1992, p. 126).

(3.71) *magasin phare, événement phare, produit phare, modèle phare* (Arnaud, 2018, p. 10)
 tienda estrella, evento estrella, producto estrella, modelo estrella

Esta equivalencia interlingüística permite reafirmar lo que Arnaud sostiene en cuanto al proceso de adjetivación, es decir que, con el tiempo, el sustantivo sufre un proceso de pauperización semántica y se convierte eventualmente en adjetivo. Este proceso también ocurre en los N2 de los compuestos de manera general, aunque en menor medida. Por ejemplo, en el compuesto “*bateau-phare*”, “*phare*” pierde el rasgo semántico “torre”, pero conserva los rasgos “marítimo” y “luz”, mientras que en los ejemplos en (3.71), solo conserva el rasgo semántico de “alta visibilidad” o, de manera más abstracta, “notoriedad”. Se puede decir lo mismo de “estrella” en español que, en los usos citados en (3.71), solo conserva el rasgo abstracto de “notoriedad” y no conserva otras relaciones con su definición original (cuerpo celeste, que corresponde a cierta forma, que emite radiación luminosa, calorífica etc.). Volviendo al francés, Arnaud estableció una lista de siete criterios para evaluar el grado de adjetivación de 29 sustantivos del francés que aparecen de manera frecuente en la posición N2 de estructuras N+N.¹²² Con la combinación de los siete criterios, el

¹²² Los criterios de Arnaud eran los siguientes:

- 1) número de coocurrencias a la izquierda: p. ej., “*magasin*”, “*valeur*”, “*événement*”, “*produit*”, “*modèle*” (“tienda”, “valor”, “evento”, “producto” y “modelo”) son coocurrencias a la izquierda de “*phare*” (lit. “faro” el equivalente en español sería “estrella”);
- 2) empleo como atributo de los verbos “*être*” (ser) o “*rester*” (quedar/seguir siendo): p. ej., el sustantivo “*clé*” (“clave/llave”) se puede encontrar en “*moment clé*” (“momento clave”) y se puede parafrasear como atributo del verbo ser en “*ce moment est clé dans l’histoire de ...*” (“este momento es clave en la historia de...”);

autor otorgó una nota a cada sustantivo según su grado de adjetivación (la nota es sobre 7). Presentamos su clasificación en la Tabla 3.2, a la cual añadimos sustantivos que cumplen una función semántica equivalente en español y cuyo uso ha sido comprobado mediante búsquedas en el *Corpus del español: Now* (Davies, 2019). Como se puede observar, la gran mayoría de los sustantivos de tendencia adjetival del francés tienen un equivalente casi literal en español, donde también aparecen como N2 en los compuestos. Un análisis sistemático del español similar al que realizó Arnaud nos permitiría corroborar si el grado de adjetivación de estos sustantivos es el mismo en las dos lenguas, pero a falta de tal análisis, las observaciones aquí realizadas permiten observar que el proceso hacia la adjetivación no parece ser exclusivo del francés. Si bien Arnaud marca una clara distinción entre lo que llama compuestos N+N y las secuencias sintácticas N+N con N2 adjetivado, en el marco de nuestro trabajo, consideramos que los límites de la categoría de los compuestos N+N son “borrosas” (Riegel, 1988a, p. 129). Más que una demarcación clara, parece existir un continuo cuyo extremo más adjetivado podría estar representado por sustantivos como “*culte*”, “*fantôme*”, etc. y otro extremo más composicional con sustantivos como “*poste*” (en “*timbre-poste*”), un N2 prototípico del fenómeno que aparece en varios trabajos de Arnaud.

-
- 3) coordinación con un adjetivo: p. ej. “*clé et fondamentale*” (“clave y fundamental”)
 - 4) modificación por un adverbio de grado: p. ej. “*très clé*” (“muy clave”)
 - 5) paráfrasis con un comparativo o superlativo: p. ej. “*de plus en plus clé*”, “*tout aussi clé que*”, “*moins clé que*” (“cada vez más clave”, “tan clave como”, “menos clave que”), etc.
 - 6) negación con “*pas*” (“no”) o “*non*” (no): p. ej. “*il n’est pas clé*” (“no es clave”), “*un aspect non clé*” (“un aspecto no clave”), etc.
 - 7) derivación con los sufijos *-té*, *-itude* o *-issime*: los ejemplos son más difíciles de encontrar con esos sufijos, tanto en francés como en español, pero Arnaud menciona “*recordité*” (“#recordicidad”), “*fantômitude*” (“#fantasmicidad”) y “*tendencissime*” (“#tendencísima”).

Tabla 3.2. Clasificación de los sustantivos de tendencia adjetival en francés (Arnaud, 2018, p. 21) con sus equivalentes en español.

No.	Nota	Sustantivo	~ en español	No.	Nota	Sustantivo	~ en español
1	6,75	<i>culte</i>	culto	16	4,58	<i>maison</i>	casa
2	6,5	<i>fantôme</i>	fantasma	17	4,5	<i>poubelle</i>	basura
3	6	<i>choc</i>	choque	18	4,07	<i>miracle</i>	milagro
	6	<i>fétiche</i>	fetiché	19	3,75	<i>pilote</i>	piloto
	6	<i>record</i>	récord	20	3,5	<i>mystère</i>	misterio
	6	<i>tendance</i>	tendencia	21	3,08	<i>béton</i>	¿?
7	5,75	<i>éclair</i>	relámpago	22	2,99	<i>spectacle</i>	espectáculo
	5,75	<i>phare</i>	estrella	23	2,91	<i>vitrine</i>	vitrina
9	5,57	<i>surprise</i>	sorpresa	24	2,74	<i>refuge</i>	refugio
10	5,25	<i>fleuve</i>	¿río?	25	2,66	<i>modèle</i>	modelo
	5,25	<i>pivot</i>	pivote	26	2,33	<i>événement</i>	¿evento?
12	5	<i>clé/clef</i>	clave	27	2,07	<i>marathon</i>	maratón
	5	<i>couperet</i>	guillotina	28	1,25	<i>cible</i>	meta
14	4,83	<i>charnière</i>	cumbre	29	1	<i>plafond</i>	techo
15	4,66	<i>plancher</i>	suelo				

Recapitulando, los compuestos N+N endocéntricos subordinativos del francés tienen su núcleo a la izquierda y presentan un alto nivel de inestabilidad en cuanto a su ortografía. Si bien el francés no se suele considerar como una lengua en la cual los compuestos N+N son productivos, ciertos patrones presentes en la lengua parecen propicios a la creación de nuevos compuestos. El nivel de productividad de los compuestos, como en español, parece estar estrechamente vinculado con las relaciones semánticas que subyacen a la unión del N1 con el N2.

3.2.2.2 Compuestos deverbales

Al igual que en español, los compuestos deverbales son muy productivos en francés y son considerados como exocéntricos y de tipo verbo-objeto (Rosenberg, 2008; Villoing, 2009). La principal diferencia formal entre los deverbales del francés y del español se encuentra en la -s genérica que se añade al final de los compuestos deverbales del español, pero no de manera sistemática en francés.

En francés, como apuntan Grevisse y Goose (2008), ha habido variaciones en la historia en cuanto a la marca de concordancia de número, pero según la ortografía recomendada vigente, los compuestos deverbales deberían seguir las mismas reglas de concordancia que cualquier sustantivo. De ese modo, si bien en español se escribe “un abrelatas/unos abrelatas”, el equivalente en francés es “*un ouvre-boite/des ouvre-boites*”. A pesar de lo que prescriben las reglas actuales del francés, Rosenberg (2007) reporta que el uso de la -s en los compuestos deverbales sigue siendo muy frecuente. Cabe subrayar que, al contrario de lo que ocurre en español, esa diferencia ortográfica no tiene repercusiones en la modalidad oral ya que la -s del plural no se pronuncia en francés (/uvrəbwat/ tanto en singular como en plural).¹²³

Otra característica del francés tiene que ver con la flexión del elemento verbal. Al contrario del español, que no muestra excepciones a este respecto, en ciertos compuestos deverbales del francés la flexión del verbo no coincide con la conjugación en tercera persona del singular. Por ejemplo, en “*soutien-gorge*” (“sostén” en español), el elemento verbal pierde la -t que lleva normalmente en su conjugación en tercera persona del singular del presente del indicativo, como lo observa Bradin (2009, p. 673).¹²⁴ Otros autores como Darmesteter (1894, p. 177), argumentan que se usa más bien la conjugación de la segunda persona del imperativo, lo cual no resuelve la irregularidad ortográfica de “*soutien-gorge*”, ya que el imperativo de “*soutenir*” es “*soutiens*”. Por su parte, Villoing (2009) considera que se trata de un tema verbal sin flexión.¹²⁵ Sea cual sea la opción privilegiada, existen excepciones en la ortografía, pero no en la pronunciación, por lo que, a pesar de las diferencias ortográficas, los deverbales del español y del francés no muestran diferencias fonológicas.

¹²³ Existen excepciones a esta falta de diferencia en la pronunciación del plural que ocurren cuando el plural del segundo elemento se forma de manera irregular. Como lo ilustra Villoing (2009, p. 5), en el compuesto atestiguado “*protège-yeux*” /prɔtɛʒzjø/ (“protege-ojos”), el segundo elemento aparece en plural. El compuesto (no atestiguado) equivalente con la palabra “ojo” en singular, “*protege-oeil*” /prɔtɛʒœj/, se pronunciaría de manera distinta.

¹²⁴ Sin embargo, dicha diferencia ortográfica no implica ninguna diferencia en la oralidad ya que ambas formas se pronuncian /sutjɛ/. Además, se podría argumentar sobre el carácter de verbal de “*soutien*” en este compuesto, ya que, si bien históricamente deriva del verbo “*soutenir*” (sostener), se trata también de un sustantivo plenamente lexicalizado que se escribe sin esa -t final.

¹²⁵ Los temas verbales representan, según Villoing (2009) una forma que carece de los rasgos morfosintácticos que llevaría un verbo dentro de un sintagma. En cuanto al verbo “*soutenir*” (“sostener”), la autora apunta que su tema verbal que entra en la composición de “*soutien-gorge*” es /sutjɛ/.

En cuanto a las relaciones semánticas dentro de los compuestos deverbales, ambas lenguas presentan patrones similares. En la mayoría de los casos, la relación entre el sustantivo y el verbo es la de paciente (3.72) y, en menor medida, un complemento circunstancial de lugar (3.73) o un agente (sujeto) (3.74). Como apunta Villoing, los ejemplos de (3.74) corresponden principalmente a topónimos y no se trata de un patrón productivo en la actualidad. Además de esas tres relaciones, que encontramos también en los deverbales del español, el francés presenta complementos de instrumento (3.75) y de intervalo de tiempo (3.76), pero son relaciones poco frecuentes.¹²⁶

- (3.72) *lèche-vitrine, ouvre-boite, gratte-papier, trotte-bébé* (Villoing, 2009, pp. 9-10)
 lame-vitrina, abre-lata, rasca-papel, trota-bebe
- (3.73) *traîne-buisson, trotte-chemin*
 anda-arbusto, trota-camino
- (3.74) *gobe-mouton, hurle-loup, pisse-vache*
 traga-oveja, grita-lobo, mea-vaca
- (3.75) *cuit-vapeur, pousse-pied*
 cuece-vapor, empuja-pié
- (3.76) *réveille-matin*
 despierta-mañana

En cuanto al referente del compuesto verbal, Villoing (2009) afirma que los referentes más comunes son los de agente e instrumento en francés. En un estudio de corpus, Rosenberg (2008) concluye que, si bien los dos referentes son productivos, los compuestos que se refieren a un instrumento son los más frecuentes. Entonces, encontramos en francés el mismo orden de productividad que observamos para el español, es decir, es decir los instrumentales en primer lugar, seguidos por los agentivos (RAE & ASALE, 2009, p. 201). Además de estos tipos de referentes, al igual que en español, los deverbales en francés también pueden designar acciones o procesos, como en (3.77).

- (3.77) *remue- méninge*
 remueve- meninges
 “lluvia/tormenta de ideas”

¹²⁶ Para estos ejemplos, solo proponemos las traducciones literales por cuestiones de concisión, dado que son compuestos que resultan difíciles de traducir al español en pocas palabras. En cuanto al ejemplo “*trotte-bébé*”, si bien “*trotter*” (“trotar”) es un verbo intransitivo y que, por lo tanto, no puede tener un complemento directo de tipo paciente, Villoing considera que “el bebé” es el paciente del verbo “*faire trotter*” (“hacer trotar”): “*Le trotte-bébé est un jouet qui fait trotter les bébés*”, “el trota bebé es un juguete que hace trotar a los bebés”.

En resumen, los compuestos deverbales del francés tienen muchos aspectos en común con los del español, tanto en cuanto a la forma como en cuanto a las relaciones semánticas entre sus componentes. La única diferencia notable que hemos apuntado concierne la -s genérica que aparece al final de los compuestos en español, pero que muestra más inestabilidad en francés. Sin embargo, se trata de una diferencia esencialmente ortográfica que no se aprecia en la modalidad oral.

3.2.3 Euskera

Como mencionamos al inicio del capítulo, el euskera presenta patrones distintos del francés y del español con respecto a los mecanismos de composición. Por su parte, el euskera comparte una serie de características con el inglés en cuanto a la formación de los compuestos, por lo que, para facilitar la comprensión, acompañaremos nuestra exposición sobre el euskera con ejemplos del inglés.

3.2.3.1 Los N+N

En euskera, al contrario de lo que hemos visto con el español y el francés, los compuestos N+N endocéntricos se forman con el núcleo a la derecha y se trata de un mecanismo de composición muy productivo (Coyos, 2004; Ortiz de Urbina, 2019). Como se puede observar en (3.78) y como comentamos en la sección 3.1.3, al contrario de lo que ocurre en una mayoría de lenguas, la posición del núcleo en los compuestos N+N endocéntricos en euskera no coincide con la posición canónica del sustantivo en un sintagma nominal.

- (3.78) a.

sagar-	arbol.a
manzana-	árbol
“manzano”	
- b.

arbola	eder.ra
árbol	bonito
“árbol bonito”	

Entonces, si bien la direccionalidad de los compuestos N+N en euskera es distinta de la del español y del francés, el orden no marcado [sustantivo + adjetivo] en el sintagma nominal es el mismo en las tres lenguas. A este respecto, Ortiz de Urbina (2019) explica que, si bien se puede considerar que el euskera es una lengua de núcleo final a nivel global, se trata de una lengua que presenta ciertas características divergentes en el sintagma nominal (p. 14). Así, al contrario de lo

que ocurre con el orden [sustantivo + adjetivo], el núcleo aparece al final en las construcciones de genitivo (3.79), en las construcciones de relativo (3.80) y en las construcciones adnominales (3.81), como en los compuestos N+N.

(3.79) Jon.en **etxe.a**

Jon.GEN **casa**.DET-SG

“La **casa** de Jon”

(3.80) ama.k irakurri du.en **liburu.a**

madre.ERG leer AUX-3SG.REL **libro**.DET-SG

“el **libro** que la madre ha leído”

(3.81) a. egur.rez.ko **mahai.a**

madera.INSTR.ADN mesa

“la **mesa** de madera”

b. Baiona.ko **best.ak**

Baiona.ADN **fiesta**.DET-PL

“las **fiestas** de Baiona”

Además, ciertos adjetivos (como los gentilicios) aparecen antepuestos al sustantivo (Trask, 2003, p. 138), aunque de Rijk (2008) no los interpreta como adjetivos sino sustantivos que forman parte de un compuesto (p. 127). Siguiendo a de Rijk, como ilustramos en (3.82), una misma palabra puede aparecer antepuesta si se considera como un sustantivo modificador dentro de un compuesto o pospuesta si se considera como un adjetivo (ver también Artiagoitia, 2002 para más detalles).

(3.82) a. erroma.tar soldadu.a

romano soldado

b. soldadu erroma.tar.ra

soldado romano

“el soldado romano”

Considerando estos ejemplos, la excepción parece ser el orden [sustantivo+adjetivo] ya que, en la mayoría de los casos, el núcleo aparece al final de los sintagmas/compuestos nominales.

En cuanto a la ortografía, según el grado de lexicalización, los dos elementos de un compuesto pueden escribirse como una sola palabra (3.83a-c), separados por un guion (3.83e), o separados por un espacio (3.83d). En los compuestos lexicalizados, ocurren procesos fonológicos que afectan al primer elemento del compuesto, como la epéntesis (3.83a), la haplología (3.83b) y la elisión (3.83c), entre otros, pero la Euskaltzaindia/Real Academia de la Lengua Vasca (1992)

diferencia de lo que hemos observado en español y en francés, muestra una mayor estabilidad en cuanto a la flexión del plural, es decir que, nunca se flexiona el primer elemento del compuesto, ni en los N+N, ni en los N+A, del mismo modo que no se flexionaría un elemento no final en un sintagma nominal.¹²⁹ Por lo tanto, esto significa que, en los compuestos N+A, el elemento que lleva las marcas flexivas no es el núcleo (el sustantivo), sino el adjetivo, como se puede observar en (3.85).

- (3.85) a. *kaska. motz.ak*
 cabeza corta.DET-PL/pelada.DET-PL
 “las cabezas rapadas”
 b. *harreman publiko.ak*
 relación publicas
 “relaciones públicas”

En cuanto a las relaciones semánticas, si bien hemos visto que los patrones de formación de compuestos N+N más productivos en español y en francés son principalmente del tipo “N2 es un tipo de o se parece a N1”, en euskera, las relaciones entre el N1 y el N2 pueden ser muy diversas. Los ejemplos de (3.86) muestran algunas de estas relaciones posibles. En todos los casos, el núcleo del compuesto aparece a la derecha. Para ilustrar el contraste entre las lenguas romances y el euskera, además de la traducción de cada ejemplo en español también aparece el equivalente en francés.

- (3.86) a. *sagar- arbola* “N2 produce N1”
 manzana árbol
 “manzano”
 “*pommier*” (en francés)
 b. *haize- eihera* “N2 funciona con el N1”
 viento molino
 “molino de viento”
 “*moulin à vent*” (en francés)
 c. *gurpil- aulki* “N2 tiene N1”

¹²⁹ Si bien es posible encontrar un sustantivo en plural seguido por un adjetivo en plural como en el ejemplo (a) en una oración, estos dos elementos no pertenecen a un mismo sintagma y por esa razón ambos se flexionan en plural.

- (a) *arbol.ak ederr.ak dira*
 árboles bonitos son
 “los árboles son bonitos”

	rueda	silla	
	“silla de ruedas”		
	“ <i>chaise roulante</i> ”		(en francés)
d.	<i>ilargi-</i>	<i>argi</i> ¹³⁰	“N2 procede de N1”
	luna	luz	
	“luz de luna”		
	“ <i>lumière de lune</i> ”		(en francés)
e.	<i>eri-</i>	<i>punta</i>	“N2 es una parte de N1”
	dedo	punta	
	“punta del dedo”		
	“ <i>bout du doigt</i> ”		(en francés)

En cuanto a la clasificación de estos ejemplos, existe el trabajo de Azkarate (1990) sobre los compuestos en euskera, comprable a los de Arnaud (2016) sobre el francés o Toquero Pérez (2018) sobre el español. Por su parte, en su gramática del euskera, de Rijk (2008) no incluye una lista exhaustiva de las relaciones semánticas de los compuestos, pero enumera 29 ejemplos que presentan diferentes relaciones semánticas para ilustrar la variedad de las relaciones posibles. Sin embargo, añade que un trabajo exhaustivo sobre el tema proporcionaría información poco relevante ya que, en varias ocasiones, una relación dada aparece en un solo compuesto (pp. 867-868). Interpretamos esto como un signo de productividad del mecanismo de composición en euskera, es decir que, no parece haber límites claros en cuanto a las relaciones semánticas posibles para formar compuestos N+N.

En resumen, el euskera presenta N+N endocéntricos con el núcleo a la derecha. La flexión de número se encuentra al final del compuesto. Las relaciones semánticas son variadas, como en inglés. Varios autores afirman que la composición N+N es un mecanismo de creación de palabras muy productivo.

¹³⁰ El lector podría preguntarse si el sustantivo “*ilargi*” (luna) es, a su vez, un compuesto de “*il+argi*”. La etimología de esta palabra ha suscitado mucho interés durante años, algunos autores parten de (1) “*ilun*” (oscuro), otros de (2) “*hil*” (matar/muerto), otros de (3) “*hil(abete)*” (mes) y, por lo tanto, podría interpretarse o bien como (1) luz oscura, (2) luz muerta o luz de los muertos, o bien como (3) luz mensual (en alusión a los ciclos de la luna). (Michelena, 1961; Uhlenbeck, 1928; Zytsar, 2000)

3.2.3.2 Los deverbales

En euskera (como en inglés), es necesario agregar un sufijo al verbo para transformarlo en sustantivo deverbal antes de crear el compuesto. En inglés, los compuestos deverbales son de tipo objeto-verbo-*er* (O+V-*er*, donde *-er* representa un sufijo de agentividad) y pueden referirse tanto a agentes, como en (3.87), como a instrumentos, como en (3.88).

(3.87) *bus* *driver*
autobús conductor
“chofer de autobús”

(3.88) *can* *opener*
lata abridor
“abrelatas”

En estos compuestos, se considera que el verbo sustantivado mediante el sufijo *-er* representa al núcleo y, por lo tanto, se consideran como compuestos endocéntricos, con el núcleo a la derecha (Fabb, 2001). Existen algunos compuestos exocéntricos de tipo verbo-objeto como “*pickpocket*” (lit. “recoge-bolsillo”, “carterista”), pero esos compuestos proceden del francés antiguo (Marchand, 1960) y el proceso para formarlos ya no es productivo en el inglés actual. En euskera, si bien Azkarate (1993) afirma que el proceso de composición V+N es productivo, los compuestos formados de ese modo no son ni equivalentes a los V+N del inglés como “*pickpocket*”, ni a los V+N del español y del francés. En ninguno de los ejemplos de V+N que aparecen en Azkarate (1993) y Hualde (2003) el sustantivo que acompaña al verbo cumple la función de objeto de este. En los V+N en euskera, el sustantivo es el núcleo del compuesto y cumple la función sintáctica de localidad, como en (3.89), o de instrumento como en (3.89).

(3.89) a. bil. toki (Azkarate, 1993, p. 225)
 reunir. lugar
 “lugar de encuentro”
 b. jos.t orratz
 coser.t aguja
 “aguja de coser”

De hecho, como comenta Hualde (2003), son compuestos subordinativos nominales con el núcleo final (p. 355), por lo que son parecidos a los N+N endocéntricos descritos en la sección anterior. Por ser diferente de los V+N en español y en francés, dejamos de lado esta estructura V+N

del euskera para concentrarnos en otra estructura más parecida a la de los N+Ver del inglés que sí es equivalente semánticamente a la estructura V+N del español (Azkarate, 1995).

3.2.3.2.1 *Los sustantivos deverbales*

En euskera, la formación de los sustantivos deverbales sigue reglas morfológicas más complejas y, por lo tanto, es preciso explorar primero la formación de sus distintos componentes antes de ver cómo se forman los compuestos deverbales (Azkarate, 1995). Los sustantivos deverbales agentivos usan los sufijos *-le* o *-tzaile*, según su base verbal y las reglas fonológicas generales de la lengua (que comportan varias excepciones).¹³¹ Cabe señalar que el único sufijo productivo para formar nuevos verbos es *-tu* y que los deverbales formados a partir de esos verbos se forman con el sufijo *-tzaile*, por lo que los derivados en *-tzaile* son más frecuentes.¹³² El ejemplo (3.90) ilustra la adaptación del préstamo “*entrenatu*”.¹³³

(3.90) <i>entrenatu</i> (verbo)	<i>entrena.tzaile</i> (sustantivo verbal)
entrenar	entren(ar).tzaile
“entrenar”	“entrenador”

Este proceso de derivación es muy productivo y se usa principalmente con verbos transitivos e inergativos, aunque se puede usar en raras ocasiones con verbos inacusativos, como en “*egoile*” (lit. “quedador”, “habitante”), “*erorle*” (lit. “caedor”, “que cae”), “*jile*” (lit. “venidor”, “que viene/llega”), “*joale*” (lit. “idor”, “emigrante”).¹³⁴ El resultado del proceso es generalmente

¹³¹ Los sufijos *-le* y *-tzaile* son dos alomorfos del mismo sufijo *-(tzaile)*, de los cuales el segundo es el más productivo (de Rijk, 2008, p. 283). Si bien las reglas fonológicas son relativamente claras, en la actualidad, existen algunos dobles como “*aurkeztzaile/aurkezle*” (“presentador”), “*motztaile/mozle*” (“esquilador”), “*ezagutztaile/ezagule*” (“conocedor”) y “*garaitztaile/garaile*” (“vencedor”) (de Rijk, 2008, pp. 283, 285).

¹³² La forma de citar los verbos en euskera corresponde a la del participio y no al infinitivo, como en español. Sin embargo, en nuestras glosas traduciremos el verbo en euskera por su equivalente infinitivo en español.

¹³³ Como muchos préstamos en euskera, identificar la fuente precisa de un préstamo romance resulta incierto. En este caso, “*entrenatu*” podría proceder directamente del francés “*entraîner*” o del español “entrenar” que a su vez procede del francés.

¹³⁴ En estos casos, todos propios del dialecto oriental (suletino), Fernández et al. (2010) proponen que el sufijo no es derivativo sino flexivo marcador de caso, por lo que no está sujeto a las mismas restricciones que su equivalente derivativo, que solo aparece con verbos transitivos e inergativos.

un sustantivo animado, pero en algunos casos puede corresponder a un adjetivo inanimado (de Rijk, 2008, p. 282), como en (3.91).

(3.91) *ireki* *irekitzaile* *soinu irekitzaile* (Elhuyar Fundazioa, 2018)
 “abrir” “abriente/abridor” “sonido abriente”

En cambio, los sustantivos deverbales que se refieren a entidades inanimadas (instrumentos) se pueden formar a partir del sufijo *-gailu* que, como palabra aislada, significa “dispositivo, aparato”, o a partir de la combinación de los morfemas *-tze* (nominalizador) y *-ko* (adnominal), que resultan en el sufijo *-t(z)eko*.¹³⁵ De este modo, se pueden usar las dos formas que aparecen en (3.92) para referirse a un instrumento que sirve para pelar. El sufijo *-t(z)eko* no tiene equivalente en español, pero se puede parafrasear como “algo/alguien que sirve para ...” (de Rijk, 2008, p. 726).¹³⁶

(3.92) *zuritu* *zuri.gailu* *zuri.tzeko* (3.92)
 pelar pel(ar).gailu pel(ar).tzeko
 “pelar” “pelador” “pelador”

¹³⁵ El caso de *-tzeko* es particular, ya que se encuentra entre la derivación (*-tze* sirve como sufijo nominalizador del verbo) y la flexión (*-ko*) es el morfema adnominal de caso adnominal.

¹³⁶ Además de los sufijos mencionados, se puede usar el sufijo *-garri*, que etimológicamente procede de “*ekarri*” (“traer”) (de Rijk, 2008, p. 335), pero la palabra resultante corresponde generalmente a un adjetivo, como en el ejemplo (a).

(a) *aberastu* *aberas.garri* (de Rijk, 2008, p. 337)
 “enriquecer” “enriquecedor”

Dado que el producto de esta deverbalización no suele aparecer en los compuestos deverbales y que, según Azkarate (1995), este sufijo es más productivo en los dialectos occidentales del euskera y nuestro estudio se lleva a cabo en la zona nororiental, no exploraremos más este sufijo. A su vez, el sufijo *-ari* (y su alomorfo *-lari*) si bien podría parecer como un sufijo formador de deverbales en algunos casos como en (b), de Rijk (2008, p. 901) apunta que este sufijo se utiliza sobre todo con los sustantivos.

(b) *kantatu* *kantatzaile* *kanta* *kantari*
 “cantar” “cantante” “canción” “cantante”

En varios casos como en (b), la forma del sustantivo y la base verbal que sirve para derivar el verbo es la misma, por lo que la diferencia de forma no es tan evidente. Además, el matiz semántico entre las dos formas, a veces, no es muy grande. En este caso, según de Rijk, “*kantatzaile*” suele referirse a cualquier persona que canta (“*singer*”), mientras que “*kantari*” suele hacer referencia a un cantante profesional (“*vocalist*”).

En resumen, los sustantivos deverbales en euskera se forman mediante la incorporación de los sufijos *-(tzai)le*, *-gailu* o *-tzeko*. Si bien pueden emplear por sí solos como los deverbales (*Ver*) del inglés, también pueden entrar en la composición de unidades léxicas más complejas como veremos a continuación.

3.2.3.2.2 *Los compuestos deverbales*

De manera general, pese a tener una morfología distinta, los sustantivos deverbales agentivos e instrumentales obedecen a las mismas reglas de composición. Como en el caso del inglés, para formar un compuesto deverbal en euskera, el objeto del verbo se coloca a la izquierda del verbo sustantivado, como en los ejemplos en (3.93).

- (3.93) a. *lursagar zuri.gailu* (con el sufijo instrumental *-gailu*)
 patata pelador
 “pelapatatas”
 b. *kale- garbi.tzaile* (con el sufijo agentivo)
 calle- limpiador
 “limpiacalles”

En cuanto a la *-s* que se usa en los compuestos del español con valor genérico, el euskera no presenta ninguna flexión equivalente en sus compuestos deverbales. Es más, es la ausencia de marca la que otorga un valor genérico al compuesto deverbal, como vemos en (3.94). Si se utiliza una marca de caso en el sustantivo, ya sea en singular o en plural, esta le da un valor no genérico al complemento y el compuesto deja de ser interpretado como tal y se vuelve un simple sintagma nominal complementado por un modificador específico (3.94).¹³⁷ Como en francés, si el referente del compuesto es plural, el compuesto se pluraliza con una marca del plural al final, como se muestra en (3.94).

- (3.94) a. *liburu- sal.tzaile.a* (Azkarate, 1995, p. 44)
 libro vendedor.DET-SG
 “el vendelibros” o “el vendedor de libros”
 b. *liburu.aren sal.tzaile.a*
 libro.GEN.SG vendedor.DET-SG
 “el vendedor del libro”
 c. *liburu.en sal.tzaile.a*
 libros.GEN.PL vendedor.DET-SG

¹³⁷ El euskera no presenta concordancia de plural interna en sus sintagmas de manera general. No sorprende, por lo tanto, que no aparezcan marcas de plural interno dentro de los compuestos.

“el vendedor de los libros”
 d. liburu- sal.tzaile.ak
 libro vendedor.DET-PL
 “los vendelibros”

Con respecto a las normas ortográficas, como en el caso de los N+N, Euskaltzaindia admite la unión de los compuestos mediante un espacio (3.93) o un guion (3.93) (1992, pp. 103-129).

En cuanto a las relaciones semánticas entre los elementos de los compuestos en euskera, el sustantivo suele cumplir la función de objeto directo del verbo, como sucede en español y en francés. Azkarate también menciona otras relaciones menos comunes, como la de localidad (3.95) y temporalidad (3.96).

- | | |
|----------------------------------|-------------------------|
| (3.95) bazter begirale | (Azkarate, 1995, p. 47) |
| rincón observador | |
| “el que observa desde un rincón” | |
| (3.96) gau jagole | |
| noche vigilante | |
| “vigilante nocturno” | |

En cuanto al referente semántico, como mencionamos al hablar de la morfología de los deverbales, los compuestos deverbales pueden corresponder de manera general a un agente animado cuando terminan en *-(tzai)le* o a un instrumento cuando terminan en *-gailu/-tzeke*. Si bien, según nuestro conocimiento, no existen estudios sobre la frecuencia relativa de estos referentes semánticos en euskera, tanto los trabajos descriptivos como los trabajos de adquisición hablan de igual manera de ambos tipos de referentes, por lo que consideraremos que ambos son igualmente productivos en el euskera actual.

3.2.4 Síntesis de la comparación de las tres lenguas

A lo largo de la sección 3.2, hemos descrito de manera detallada y sistemática los compuestos nominales N+N y deverbales en español, francés y euskera. A partir de esta descripción detallada, sintetizamos ahora las principales diferencias y semejanzas entre las tres lenguas del estudio.

En cuanto a los N+N, el francés y el español comparten la misma direccionalidad, es decir, tienen un núcleo inicial, mientras que el euskera presenta un orden con el núcleo al final. La flexión

del plural en los compuestos N+N en español y en francés suele aparecer solamente en el núcleo, pero el francés exhibe más inestabilidad al respecto y, en ciertas situaciones, presenta una marca en ambos componentes. En cambio, el euskera muestra más estabilidad y presenta siempre una marca de plural al final del compuesto. Si bien la composición N+N es poco productiva en español y en francés, la gran variedad de relaciones semánticas encontradas en los compuestos atestiguados parece indicar que se trata de una opción posible para acuñar nuevas palabras, sobre todo cuando la relación entre el núcleo y su modificador se asemeja a la relación que existe entre un sustantivo y un adjetivo restrictivo. En cambio, la composición N+N es muy productiva en euskera y las relaciones semánticas entre los componentes parecen más variadas.

En cuanto a los compuestos deverbales, el español y el francés también presentan la misma direccionalidad, con el verbo al inicio, mientras que el euskera presenta el patrón opuesto, con un elemento verbal al final. El elemento verbal aparece en tercera persona del singular en español y en francés, mientras que en euskera aparece con un sufijo agentivo o instrumental. Los compuestos deverbales del español suelen presentar una -s final genérica, marca que también aparece en francés, pero de forma menos sistemática. De hecho, las reglas prescriptivas del francés recomiendan evitar la -s genérica y hacer la concordancia como se haría con cualquier sustantivo. En euskera, los compuestos deverbales también concuerdan en número como cualquier sustantivo, con una marca del plural sobre el determinante, que aparece afijado al final del compuesto. A modo de resumen, la Tabla 3.3 presenta los mismos ejemplos en las tres lenguas, con los núcleos en negrita y las marcas de plural subrayadas.

Tabla 3.3. Ejemplos de compuestos nominales en las tres lenguas del estudio.

Tipo de compuesto	Español	Francés	Euskera
N+N	el hombre rana	l' homme -grenouille	igel gizona
animado	los hombres rana	les hommes -grenouilles	igel gizonak
N+N	el barco pirata	le bateau pirate	pirata itsasonztia
inanimado	los barcos pirata	les bateaux pirates	pirata istasontziak
Deverbal	el limpiabotas	le lave -botte	zapata- garbitzalea
animado	los limpiabotas	les lave -bottes	zapata- garbiltzaleak
Deverbal	el lavaplatos	le lave -vaisselle	baxera- garbigailua
inanimado	los lavaplatos	les lave -vaisselles	baxera- garbigailuak

Para cada tipo de compuesto, mostramos un ejemplo animado e inanimado. Si bien solo hemos abordado el contraste entre animado e inanimado al hablar de los compuestos deverbales, presentamos también N+N que corresponden al par animado e inanimado para completar la Tabla 3.3 y profundizaremos este contraste al hablar de los estudios de adquisición en la sección 3.4.

3.3 Análisis morfosintáctico y semántico

Desde un punto de vista general, el fenómeno de composición es universal, pero lo que subyace a la formación de los compuestos sigue ciertas restricciones que suelen variar de una lengua a otra. Por ejemplo, como hemos comentado en la sección anterior, el euskera favorece los N+N con un núcleo final, pero en francés y en español predominan los N+N con un núcleo inicial.

En la presente sección, exploraremos desde diferentes perspectivas cómo se puede explicar el fenómeno de la composición, principalmente en español, pero siempre comparándolo con las demás lenguas de nuestro estudio y con el inglés, que nos servirá de punto de referencia, dado el gran número de estudios existentes sobre el fenómeno en esta lengua. Los análisis que se presentarán a continuación se interesan por los compuestos desde la sintaxis, la morfología y la semántica. Al presentar estos diferentes niveles de análisis, buscamos exponer cómo cada uno de esos niveles es necesario para explicar los procesos que subyacen y restringen la formación de los compuestos en las tres lenguas del estudio. Si bien los compuestos N+N y los compuestos

deverbales comparten ciertas características, los presentaremos de manera separada en las subsecciones 3.3.1 y 3.3.2, respectivamente.

3.3.1 Compuestos N+N

Independientemente del marco teórico, una explicación satisfactoria sobre la formación de los compuestos N+N endocéntricos debe aclarar por qué cierto orden (núcleo inicial vs. final) predomina en una lengua dada y por qué ciertos sustantivos se pueden unir para formar un compuesto, pero otros no. En el programa minimalista (Chomsky, 1995), del marco generativo, se remite la primera parte del problema a las reglas formales que interactúan con ciertos rasgos de los elementos del lexicon y se tiende a olvidar el aspecto semántico (Lieber, 2016). En los enfoques basados en el uso, se contempla el problema mediante la noción de construcción, que une a la vez la forma y el significado. A continuación, presentaremos algunas propuestas generativas que luego contrastaremos con propuestas basadas en el uso.

3.3.1.1 El parámetro de la composición

Una propuesta para analizar la estructura de los compuestos N+N en español se basa en el parámetro de la composición (TCP; *The Compounding Parameter*), introducido por Snyder (1995). Según este parámetro, las diferentes clases de palabras de una lengua pueden llevar un rasgo [\pm afijable] y, por consiguiente, se dice que esas lenguas tienen el parámetro [\pm TCP]. Los sustantivos del español llevan el rasgo [-afijable], que impide la afijación de un sustantivo a la izquierda del núcleo y favorece alternativas que siguen las reglas sintácticas de la lengua en cuestión (p. ej., el sintagma nominal complementado por un sintagma preposicional en español y en francés).

Siguiendo la tradición generativa, Snyder (2016b) vincula varios fenómenos de la lengua con el TCP. Además de la creatividad de los compuestos N+N, otros fenómenos relacionados con el parámetro son las partículas separables y las cláusulas adjetivas resultativas, que ilustramos en los ejemplos (3.97) y (3.98), respectivamente.

- (3.97) a. *run up the bill* (Liceras et al., 2002, p. 5)
 correr arriba.PARTÍCULA la cuenta
 b. *run the bill up*
 correr la cuenta arriba.PARTÍCULA
 “subir la cuenta”

(3.98) *wipe the table clean*
 limpiar la mesa limpia
 “limpiar la mesa”

(Snyder, 2016b, p. 126)

El ejemplo (3.97) ilustra que, si bien la partícula “*up*” puede ir contigua al verbo, también puede ir separada. No hay equivalentes en español, en francés o en euskera de este tipo de estructuras, por lo que se considera que esas lenguas no permiten las partículas separables. A su vez, el ejemplo (3.98) ilustra que se pueden construir cláusulas adjetivas resultativas en inglés, pero que no existen en francés, en español o en euskera. En la Tabla 3.4, presentamos un resumen de la relación entre esos fenómenos y las lenguas de nuestro estudio, según la propuesta de Snyder.¹³⁸

Tabla 3.4. Fenómenos vinculados con el TCP y su parametrización según Snyder (2012, p. 287).

Lengua	TCP	Partículas separables	Adjetivas resultativas	Composición N+N creativa
Inglés (lenguas germánicas)	+	Sí	Sí	Sí
Español y francés (lenguas romances)	-	No	No	No
Euskera	-	No	No	No

Un problema con esta tabla concierne notablemente la noción de creatividad de los compuestos N+N, la cual hemos abordado en la sección 3.1.7. Si bien hemos mencionado en la sección anterior que el francés y el español permiten cierta creatividad, la posible creación de compuesto se hace siempre con el núcleo a la izquierda, lo que Snyder no considera como verdaderos compuestos N+N escuetos (*bare-stem compounds*). Por su parte, en euskera sí que

¹³⁸ Snyder se basa en un sondeo que cuenta con informantes de una variedad de lenguas de diferentes grupos: afroasiáticas (árabe egipcio, hebreo), austroasiáticas (jemer), austronesias (javanés), ugrofinesas (húngaro), indo-europeas (germánicas: inglés, alemán; romances: francés, español; eslavas: ruso, serbo-croata), Niger-Congo (bantú: lingala), y sino-tibetanas (chino-mandarín), más el japonés, el coreano, la lengua de los signos americana y el euskera (Snyder, 2012, p. 287). Todas coinciden en cuanto las adjetivas resultativas y la composición creativa, solo un subconjunto de las [+Adjetivas resultativas] y [+Composición creativa] tienen la configuración [+partículas separables].

existen compuestos N+N con el núcleo a la derecha como en inglés, pero, a pesar de ello, Snyder clasifica el euskera como una lengua [-TCP].¹³⁹ Para explicar esa clasificación, Snyder argumenta que, en los compuestos N+N del euskera, el N1, que actúa como modificador del N2, no aparece en su forma escueta sino en su forma “modificacional”, es decir, con una terminación diferente (2016b, p. 95). En el ejemplo proporcionado por Snyder, que reproducimos en (3.99), “*giza*” (“humano”) representa la forma modificacional de “*gizon*” (“hombre”).

(3.99) *giza* *hizkuntza*¹⁴⁰
 humano lengua
 “lengua humana/lenguaje humano”

Del mismo modo que Snyder (2016b, p. 95) intuye, a partir de de Rijk (2008, p. 854), que “*giza*” es la forma modificacional de “*gizon*”, Artiagoitia et al. (2016) afirman que la primera es la forma “composicional” de la segunda. Sin embargo, de Rijk (2008, p. 854) afirma que, cuando un compuesto se refiere exclusivamente a algo del “hombre” y no de la “humanidad”, se usa “*gizon*”, como en el ejemplo (3.100).¹⁴¹

(3.100) *gizon- dantza*
 hombre danza
 “danza de hombre” (el nombre de una danza específica)

Considerando esto, cabe preguntarse si los cambios morfofonológicos que afectan al elemento no-nuclear en ciertos compuestos representan realmente un paso necesario para formar compuestos N+N en euskera, como afirma Snyder, o si son más bien el resultado de un proceso de lexicalización. Hualde (2003, p. 354) menciona que esos cambios suelen ocurrir más a menudo en los compuestos lexicalizados desde hace cierto tiempo (“*old compounds*”). Además, la mayoría de

¹³⁹ Snyder clasificó inicialmente el euskera en lengua [+TCP] (Snyder, 2001, pp. 329-330), para luego cambiar su clasificación a [-TCP] (Snyder, 2012, p. 287).

¹⁴⁰ Snyder escribe “*giza kuntza*” y lo traduce por “*human language*”, pero podemos imaginarnos que el autor quiso escribir “*hizkuntza*” que efectivamente significa “lengua” y que aparece en de Rijk (2008, p. 854), la obra de referencia citada por Snyder. Si bien “*gizakuntza*” en una sola palabra aparece en el diccionario *Elhuyar* (Elhuyar Fundazioa, 2018), tiene una interpretación distinta, es decir, “hominización”.

¹⁴¹ El ejemplo (3.100) no es de de Rijk (2008) porque los ejemplos proporcionados por el autor eran compuestos lexicalizados en los que era difícil identificar cada componente. Nuestro ejemplo procede del diccionario *Elhuyar* (Elhuyar Fundazioa, 2018)

los cambios fonológicos descritos por de Rijk (2008, pp. 854-857) y que Snyder considera necesarios para formar los compuestos no son propios de la formación de los compuestos, sino que obedecen a las reglas fonológicas generales del euskera. Es más, esas reglas fonológicas no ocurren de manera sistemática para un mismo sustantivo y, cuando se aplican, lo hacen en compuestos ya altamente lexicalizados. Entonces, más que un requisito para acuñar nuevos compuestos, consideramos que la forma modificacional mencionada por Snyder representa simplemente un paso hacia la posterior lexicalización de estos compuestos. Por ejemplo, de Rijk explica que ciertas palabras acabadas en -r pierden esa -r final cuando aparecen ante una consonante. Reproducimos en (3.101) el ejemplo que presenta de Rijk (un compuesto N+A lexicalizado) formado a partir de “lur” (tierra), que comparamos con un N+N lexicalizado en (3.102) y otro N+N que muestra un menor grado de lexicalización en (3.103).

- (3.101) lu. berri (de Rijk, 2008, p. 854)
 tierra-nuevo
 “tierra noval”
- (3.102) lur. sagar (Elhuyar Fundazioa, 2018)
 tierra manzana
 “patata”
- (3.103) lur- esfera
 tierra esfera
 “esfera terrestre”

En los ejemplos (3.102) y (3.103), se puede notar que el primer elemento del compuesto aparece de forma inalterada y escueta (sin morfema derivativo, sin aposición), como lo hacen los modificadores los compuestos con un menor grado de lexicalización en el euskera actual, siguiendo las recomendaciones de la Euskaltzaindia (1992). Además, estos ejemplos nos recuerdan que el euskera presenta compuestos N+N que poseen un núcleo final y que permiten cierta creatividad en su interpretación. Considerando esto, sería pertinente recategorizar el euskera con el rasgo [+composición creativa], en contra de la propuesta de Snyder, pero el euskera sería la única excepción que permite la composición creativa sin permitir las oraciones adjetivas resultativas. Entonces, el macro-parámetro pasaría a ser un parámetro únicamente aplicable a la composición creativa, de manera independiente de las demás propiedades mencionadas por Snyder, al menos en el caso del euskera. Ahora volviendo al francés y al español, a parte de categorizar esas lenguas como lenguas en las que no se pueden formar compuestos, el TCP no nos explica por qué, en estas

lenguas, ciertos compuestos N+N con núcleo inicial son posibles y, hasta cierto punto, creativos, como hemos visto en la sección 3.1.7.¹⁴²

Por otro lado, como menciona Liceras et al. (2002, p. 5), siguiendo los principios de la lingüística generativa, esta propuesta también supone que la formación de compuestos mediante la afijación de un N a otro N representa la opción marcada, una opción que no deberían preferir los aprendientes de L2. Sin embargo, los resultados del estudio de Liceras et al. (2002) con aprendientes de español como L2 demuestran que la composición de tipo N+N representa una estrategia presente en la interlengua de sus participantes de EL2, en contra de la hipótesis de Snyder.

3.3.1.2 El marcador de palabra y la restricción de los corchetes dobles

Una propuesta teórica alternativa presentada en Liceras et al. (2002) se basa en la noción del marcador de palabra (MP, *Word Marker*), introducida por Harris (1991) para explicar el género en español y aplicada a la formación de compuestos por Piera (1995). Según esta hipótesis, todos los sustantivos de una lengua como el español poseen un MP, mientras que los sustantivos de una lengua como el inglés no lo tienen. Siguiendo los ejemplos de Liceras et al. (2002, pp. 5-7), en la sintaxis, el MP genera un par de corchetes adicionales alrededor del sustantivo. Los ejemplos (3.104) y (3.105) representan respectivamente un sustantivo en inglés y su equivalente en español.

(3.104) [dog]

(3.105) [[perr-]MP]

Según la restricción de los corchetes dobles (*Double Bracket Restriction*; Piera, 1995), es imposible insertar un sustantivo a la izquierda de unos dobles corchetes, por lo que, si bien es posible adjuntar “*police*” a la izquierda de “*dog*” como en (3.106), resulta agramatical la adjunción de “*policía*” a la izquierda de “*perro*” como en (3.107).¹⁴³ En cambio, la proyección del elemento

¹⁴² En una nota de pie de página Snyder (2016a), solo alude a la creatividad de los compuestos V+N en las lenguas romances: “While the Romance languages are clearly [-TCP], there do seem to be a few types of compounding that are creative, and these merit greater attention from acquisitionists. These include the exocentric V-N compounds to name instruments and (occasionally) agents ...” (p. 110)

¹⁴³ Esta adjunción del segundo elemento ocurre como una operación morfológica pre-sintáctica según Piera. Liceras no entra en los detalles, pero propone como alternativa la posibilidad de que esa operación de adjunción ocurra dentro del léxico, antes de que el compuesto entre en la numeración (Liceras et al., 2002, p. 27)

inicial “perro” sí es posible, ya que, a su derecha solo se encuentra un corchete simple, como se puede observar en el ejemplo (3.108). Esta restricción es también la que impide la recursividad de los compuestos en español, pero no en inglés. De la misma manera, el compuesto en inglés solo tiene un corchete a su izquierda y permite la adjunción de otro sustantivo, pero el compuesto en español tiene corchetes dobles a su derecha (donde se tendría que adjuntar el sustantivo), lo que bloquea la adjunción, como muestran los ejemplos (3.109) en inglés y (3.110) en español.

- (3.106) $N[\text{police} [\text{dog}]]$
 (3.107) $*N[[[\text{policí-}]MP] [[\text{perr-}]MP]]$
 (3.108) $N[[[\text{perr-}]MP] [[\text{policí-}]MP]]$
 (3.109) $N[\text{pet} [\text{police} [\text{dog}]]]$
 (3.110) $*N[[[\text{perro-}]MP] [[\text{policí-}]MP]] [\text{mascot-}]MP]]$

Si bien nos podemos apoyar en esta misma hipótesis para explicar la direccionalidad de los compuestos N+N en euskera –sin MP como el inglés– no queda claro si los sustantivos del francés tienen o no un MP como los sustantivos del español. En un trabajo sobre el género gramatical, Meisel (2018, p. 660) usa la expresión de Harris (1991) para aludir a las terminaciones de los sustantivos francés.¹⁴⁴ Si bien Meisel usa el término solo para referirse a la terminación (fonológica) de las palabras y no detalla en qué estas terminaciones cumplen el mismo papel que los MP del español según Harris, nos parece adecuado asumir que los sustantivos del francés también tienen un MP, pero con una forma más variable que en español. Más que un marcador con una forma definida, entendemos el MP como una capa morfológica que tienen las lenguas romances y que permite ciertas operaciones morfosintácticas (p. ej., la concordancia de género) que no se aplican a otras lenguas. Siendo así, el simple hecho de que ambas lenguas presenten un MP, sea cual sea su forma, nos permite aplicar el mismo análisis que Liceras et al. (2002) y Piera (1995) han aplicado al español. Ahora bien, dicha propuesta permite explicar de manera satisfactoria la diferencia de direccionalidad entre las dos lenguas romances y el euskera, pero no toma en cuenta las restricciones semánticas que parecen aplicarse al francés y al español, pero no al euskera. A continuación, exploramos consideraciones semánticas que, según lo presentado en la sección 3.2, contribuyen a restringir la composición nominal de manera variable de una lengua a otra.

¹⁴⁴ Kayne (2003) también habla de los MP en francés en términos compatibles con las explicaciones de Harris (1991), pero solo alude a los MP en los clíticos, no en los sustantivos.

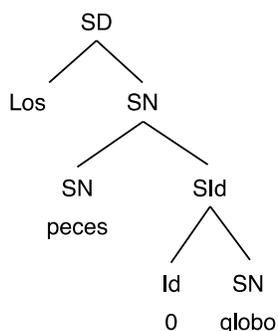
3.3.1.3 Consideraciones semánticas

Si bien las propuestas centradas en la sintaxis que hemos revisado hasta aquí proponen mecanismos para explicar la direccionalidad de los compuestos, estas no ofrecen explicaciones sobre las restricciones en las relaciones semánticas posibles entre sus componentes. Por ejemplo, la propuesta basada en el marcador de palabra nos permite descartar la creación de compuestos con un núcleo final en español, pero no nos permite explicar por qué un compuesto como “perro policía” (espejo de “*police dog*” en inglés) es posible y comprensible por un hablante nativo de español, mientras que “comida mar” (espejo de “*seafood*”) resultaría incomprensible o, por lo menos, extraño. Hemos visto en la sección 3.2 que, si bien los compuestos N+N del español y del francés presentan una variedad de relaciones semánticas, existen pocas relaciones muy productivas. Marqueta Gracia (2017) considera que las relaciones semánticas posibles de los compuestos del español se resumen en una sola función, la de “identificación”, que se parafrasea como “N1 ... como (un) N2”, y propone los ejemplos que reproducimos en (3.111).

- (3.111) a. Ese pez es redondo **como** un globo (pez globo).
b. Esa ciudad se usa **como** dormitorio (ciudad dormitorio).
c. Esa actriz es brillante **como** una estrella (actriz estrella).
d. Ese papel es rugoso **como** el cartón (papel cartón).
e. Ese actor trabaja también **como** director (actor-director).

Dicha función de identificación descarta compuestos como “*comida mar”, que no son aceptables en español, ya que no se pueden parafrasear como “*esa comida es como el mar”. Basándose en esta única función para formar compuestos en español, la autora postula la estructura sintáctica que reproducimos en (3.112) y que consta de una proyección funcional *ad hoc* (Sintagma Identificacional, SId) entre los dos elementos del compuesto.

(3.112) (Marqueta Gracia, 2017, p. 14)



Este nodo intermedio sería el responsable de la restricción semántica, es decir que solo permitiría la inserción de un sustantivo que corresponda a la función de identificación postulada por la autora. En cambio, los compuestos del inglés, que no tienen este SId, no presentan esta restricción semántica. Para explicar otras relaciones que no parecen cumplir con la función de identificación, como “jamón York”, la autora explica que éstos no corresponden a la estructura presentada en (3.112), sino que se tienen que analizar como sintagmas en los que la preposición “de” es omitida.¹⁴⁵

Si bien esta propuesta permite explicar la restricción semántica de los compuestos en español hasta cierto punto (y en francés si consideramos que esta lengua también tiene el SId), los límites de esta restricción no parecen adecuados para todos los compuestos. De las diversas relaciones semánticas en los compuestos del francés y del español que hemos visto en la sección 3.2, algunas de ellas no se pueden explicar ni mediante la función de identificación, ni mediante la elisión de la preposición “de” (o de otras preposiciones). Según Arnaud (2016), los compuestos N+N endocéntricos relacionales no implican únicamente una relación identificacional (“como”), sino una variedad de relaciones distintas. Por ejemplo, para explicar compuestos como “*imprimante laser*” (“impresora láser”), Arnaud presenta una relación que parafrasea como “el no-núcleo es el fenómeno/proceso/modo de acción/fuente de energía/ambiente que permite al núcleo funcionar o existir”. Esta relación subyace a la formación de 59 compuestos del corpus de Arnaud (2016) y, por lo tanto, no se puede tratar como una mera excepción. Según la propuesta de Marqueta Gracia, si un compuesto no se puede analizar mediante la función “identificacional”, la otra opción consiste en la preposición elidida (“impresora de láser”), opción que resulta parcialmente satisfactoria en este caso. Si bien podríamos postular la existencia de una preposición elidida en algunos compuestos formados con “láser” en posición de N2, se puede observar en la Tabla 3.5 que la opción sin preposición es de uso más frecuente en todos los casos, por lo que resulta difícil explicar la elisión de una preposición que raramente aparece en estos contextos.¹⁴⁶

¹⁴⁵ La idea de reinterpretar los compuestos como sintagmas en los que se elide la preposición no es nueva, Arnaud (2003, 2016) atribuye esta hipótesis a Grieve-Schumacher (1960) para los compuestos del francés.

¹⁴⁶ Los datos de la Tabla 3.5 proceden de nuestras búsquedas de “láser” en el *Corpus del Español: Now* (Davies, 2019). Hemos limitado el listado a las quince coocurrencias más frecuentes de “N1(s) + láser” y hemos comprobado las variantes con todas las preposiciones posibles y reportado solo las preposiciones que tenían por lo menos una ocurrencia (con, de, por, mediante, a, en).

Tabla 3.5. Frecuencia de los compuestos cuyo N2 es "láser" y sus variantes con preposiciones.

	Compuesto sin preposición	Variantes con preposición
1	rayó(s) láser (2019)	de (9)
2	tecnología(s) láser (945)	de (45), en (4)
3	depilación láser (718)	con (76), a (7), por (5), de (2)
4	sable(s) láser (966)	de (5)
5	luz/luces láser (898)	de (40), con (1)
6	espada(s) láser (685)	de (5)
7	puntero(s) láser (636)	de (10), a (1)
8	tratamiento(s) láser (366)	con (289), de (54), mediante (2), por (3), a (1)
9	impresora(s) láser (513)	de (6), a (2)
10	arma(s) láser (393)	de (9), con (3), por (1)
11	(micro)cirugía(s) láser (379)	con (63), de (8), por (4), mediante (1)
12	escáner(es) láser (328)	de (17), con (15), por (1), mediante (1)
13	sistema(s) láser (301)	de (66), con (1)
14	corte(s) láser (204)	por (46), con (27), en (13), a (9), de (6), mediante (1)
15	mira(s) láser (204)	de (2)

Cabe subrayar que la mayoría de estos compuestos son de creación reciente, ya que se basan en una tecnología relativamente nueva. Además, en todos los casos existe un compuesto equivalente en inglés (con la direccionalidad opuesta), por lo que se podría afirmar que son simplemente calcos del inglés. Sin embargo, más que simples calcos del inglés, la gran variedad de combinaciones nos lleva a considerar que se trata de un proceso de composición productivo del español, particularmente utilizado para integrar términos tecnológicos extranjeros en los compuestos (p. ej., “radar”, “diesel”).

Desde una perspectiva más en línea con los acercamientos basados en el uso, Arnaud (2016) analiza la estructura de los compuestos siguiendo los principios de la lingüística cognitiva (Langacker, 1987; Ryder, 1994; Booij, 2009, 2010; Masini, 2009; Jackendoff, 2011, 2016). Si bien los diferentes autores presentan sistemas con sus particularidades, la noción de esquema o construcción (términos que utilizaremos de manera intercambiable) es el punto común entre todos.¹⁴⁷

Desde esta perspectiva, los compuestos nominales son considerados como construcciones que se memorizan como cualquier otro elemento léxico. Los nuevos compuestos se crean por analogía con los compuestos existentes, que actúan como modelos (*linguistic templates*; Ryder, 1994, pp. 79-80). La frecuencia también tiene un papel importante, ya que los modelos más frecuentes son más susceptibles de fomentar la creación de nuevos compuestos. Por ejemplo, si un esquema [N1+N2, N1 es como N2] corresponde a varios compuestos en español, es más probable que se acuñen compuestos siguiendo este modelo. En las gramáticas de construcción, para representar los compuestos, existen esquemas de diferentes niveles de abstracción relacionados entre sí en los que se especifican diferentes propiedades relacionadas con la forma y el significado. Los detalles de las propiedades incluidas en las construcciones varían de una teoría a otra, pero en todos los casos se organizan en torno a estos dos ejes. Por ejemplo, la construcción según Croft (2001) es bipartita: en la parte “forma” se ubican las propiedades sintácticas, morfológicas y fonológicas, mientras que las propiedades semánticas, pragmáticas y discursivo-funcionales se encuentran en la parte “significado” (p. 18). En lo que sigue, por razones de concisión, nos concentraremos en la forma y el significado, sin entrar en los detalles técnicos de las diferentes gramáticas de construcciones.

Partiendo de un nivel muy abstracto, los compuestos se pueden representar mediante una construcción como [X+Y] en la cual ni la categoría de los componentes, ni la direccionalidad del compuesto están explicitadas. Dado que este esquema permite combinar dos elementos cualesquiera, es necesario que de este nivel de abstracción se desprendan esquemas más concretos. Entonces, en otro nivel encontramos esquemas más específicos como [N+N] y luego, como [N1+N2, N1 protege de N2] en el caso del francés o [N1+N2, N2 protege de N1] en el caso del

¹⁴⁷ Referimos al lector a la sección 1.2.1.2 del capítulo 1 sobre los detalles de la noción de construcción, pero repetimos a continuación los elementos esenciales para entender cómo podemos aplicar dicha noción a los compuestos.

	b. <i>*chaise roue</i>	
	c. <i>wheelchair</i>	N2 tiene N1
	d. <i>gurpil-aulki</i>	
(3.117)	a. <i>*barco vapor</i>	N1 es propulsado por N2
	b. <i>*bateau vapeur</i>	
	c. <i>steamboat</i>	N2 es propulsado por N1
	d. <i>lurrun.ontzi</i>	
(3.118)	a. <i>*dedo punta</i>	N1 es una parte de N2
	b. <i>*doigt bout</i>	
	c. <i>finger tip</i>	N2 es una parte de N1
	d. <i>herri.punta</i>	

Esta lista no es exhaustiva y podría alargarse de manera considerable si incluyéramos todas las posibilidades del inglés (Clark et al., 1985, p. 84) o del euskera (de Rijk, 2008), pero resulta suficiente para ilustrar que cualquier hablante del inglés o del euskera está expuesto a una mayor variedad de compuestos nominales que un hablante de español o de francés. Dado que la frecuencia y la analogía son dos claves para la creación de los esquemas, a medida que un hablante de lenguas como el inglés o el euskera vaya adquiriendo nuevos compuestos, ampliará su repertorio de esquemas que corresponden a la forma [N+N]. Por consiguiente, es más probable que ese hablante produzca compuestos [N+N] que un hablante de español, que tiene acceso a una menor variedad de esquemas [N+N] y una frecuencia menor de esas construcciones en su *input*. La existencia o no de un esquema concreto en una lengua dada es precisamente lo que condiciona su producción (Heyvaert, 2009). Por lo tanto, ya que en el español actual no existen ejemplares como “*finger tip*”, es poco probable que un hablante nativo produzca “dedopunta” (aunque tenga la direccionalidad adecuada). Desde una perspectiva basada en el uso, las construcciones que subyacen a la formación de los compuestos no solo varían de una lengua a otra, sino que también pueden variar de un hablante a otro. Dado que cada hablante modela sus construcciones en función del *input* que recibe, es posible que dos hablantes de una misma lengua se representen un mismo compuesto bajo dos construcciones distintas. Por ejemplo, para un hablante que nunca ha sido expuesto a los ejemplos en (3.113), pero que sí ha oído compuestos como “*assurance chômage*” (lit. “seguro desempleo”, “seguro de desempleo”) y “*assurance vie*” (lit. “seguro vida”, “seguro de vida”), es posible que la construcción [N1+N2, N1 protege de N2] no forme parte de su repertorio léxico. En cambio, es posible que tenga una construcción más concreta como [*assurance*+N2, *assurance* tiene relación con N2].

A su vez, la existencia de una construcción no impide la existencia de varias construcciones de forma diferente con un significado equivalente, ya que la construcción corresponde a un par forma-significado. Así, a pesar de la posible existencia de una construcción [N1+N2, N1 protege de N2], es posible que exista otro esquema [N1+Prep+Det+N2, N1 protege de N2] para producir sintagmas como el equivalente de “*cough syrup*”, que ilustramos en (3.119). Entonces, para ilustrar una relación similar, es posible que, por analogía, se formen tanto compuestos que siguen los ejemplos de (3.113) como sintagmas según el modelo en (3.119).¹⁴⁸

(3.119) *sirop pour la toux*
 jarabe para la tos

Resulta interesante subrayar que, tanto en francés como en español, el uso de la preposición “*pour/para*” no es mucho más transparente que la ausencia de cualquier elemento en el equivalente en inglés o en euskera (“*eztul ziropo*”, lit. “tos jarabe”), es decir, que el uso de la preposición “para” no indica de manera precisa si dicho jarabe sirve “para (curar)” o “para (fomentar)” la tos. Este uso poco previsible de la preposición nos permite volver sobre los compuestos formados con el sustantivo “láser” que presentamos en las páginas anteriores. Ahora usando el concepto de construcción para explicar compuestos como “tratamiento laser”, podríamos postular que existe en español un esquema [N1+N2, N1 funciona con la tecnología de N2]. Como muestran los datos que presentamos en la Tabla 3.5, este esquema compite con otros esquemas prepositivos que muestran una frecuencia de uso menor. Las razones que podrían explicar que ciertos hablantes prefieren un esquema prepositivo versus un esquema sin preposición son varias, pero queremos insistir en un pilar del acercamiento basado en el uso: la frecuencia. Si un hablante dado ha oído principalmente expresiones como “tratamiento con antibióticos”, “medicamentos”, “células”, “plasma”, “quimioterapia”, etc., es probable que tenga una construcción como [N1+con+N2, N1 se cura con la ayuda de N2] y recurra a esta a la hora de producir “tratamiento con láser”. En cambio, si el *input* recibido por otro hablante le ha permitido crear un esquema del tipo [N1+N2, N1 funciona con la tecnología de N2], es posible que este se use a la hora de producir “tratamiento láser”. Resulta difícil saber si el esquema en cuestión es el calco del esquema correspondiente del inglés o si se ha creado simplemente por la posible confusión entre las diferentes preposiciones (o una

¹⁴⁸ Es posible que la diferencia entre las dos construcciones se encuentre en el rasgo [+/- abstracto] que parece ser una de las diferencias entre una estructura y la otra.

combinación de los dos factores), pero la existencia de un número considerable de compuestos que siguen este modelo en español nos obliga a aceptar la existencia de un tal esquema, sea propio del español o no.

Este último punto es de particular interés en el marco de nuestro estudio, ya que exploramos la influencia interlingüística. Si un hablante tiene ciertas construcciones en su L1, a la hora de adquirir una L2, la creación de nuevas construcciones no estará solamente constreñida por el *input* de esta lengua, sino también por las construcciones de su L1. Luego, a la hora de adquirir una L3, las construcciones disponibles podrían proceder tanto de la L1 como de la L2. No todos los autores están de acuerdo en cuanto al nivel de integración de las construcciones en las distintas lenguas de un hablante (ver, por ejemplo, Hilpert & Östman, 2014), si bien parece claro que todas estas construcciones interactúan entre sí.

Resumiendo, en la presente sección, hemos presentado análisis basados en la gramática generativa que permiten sobre todo explicar la direccionalidad de los compuestos N+N de una lengua a otra, postulando parámetros más generales de la lengua, o categorías funcionales que restringen las categorías léxicas de cada lengua. Posteriormente, hemos presentado análisis más afines a las perspectivas basadas en el uso, que usan la noción de construcción para explicar las restricciones semánticas que pueden variar de una lengua a otra, postulando diferentes construcciones para cada lengua. Ambas equivalentes en cuanto a la direccionalidad, la explicación basada en la noción de construcción proporciona más detalles sobre los matices semánticos entre una lengua y otra.

Volveremos sobre las predicciones que podemos establecer a partir de estos análisis y la interacción interlingüística en la sección 3.6.1. A continuación, exploramos las propuestas de análisis morfosintácticos y semánticos de los compuestos deverbales.

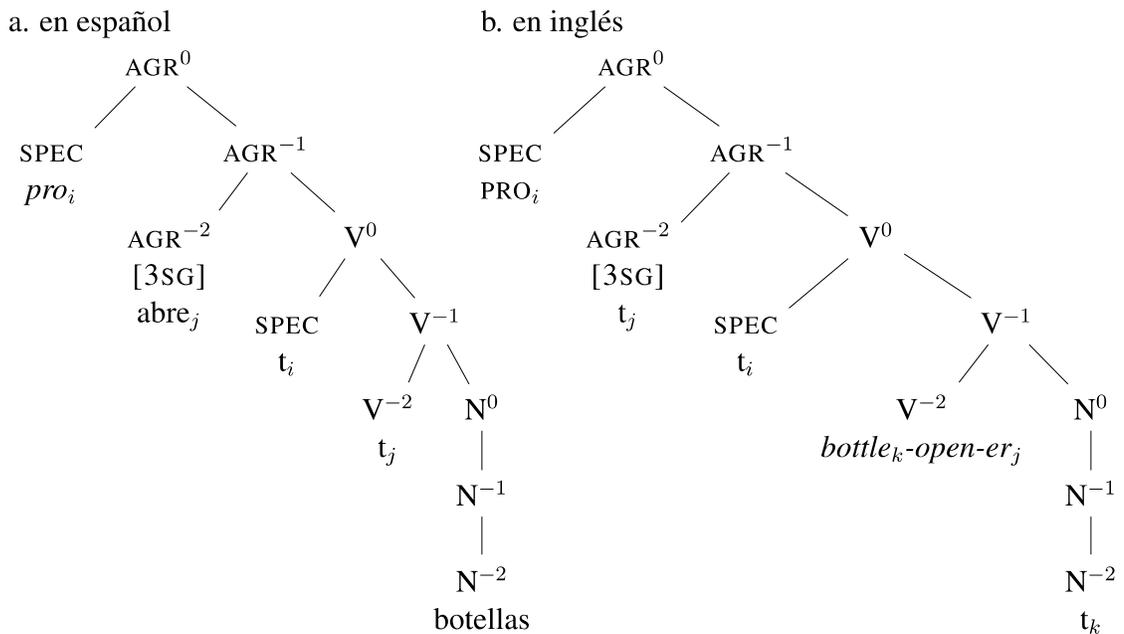
3.3.2 Compuestos deverbales

Los compuestos deverbales, como los compuestos N+N, constan de dos elementos que tienen relaciones comparables a las que podríamos encontrar dentro de diferentes sintagmas que no son considerados como compuestos. En el caso de los compuestos deverbales, esta relación es la de un verbo conjugado con un argumento interno (generalmente el tema). A continuación, para dar cuenta de la forma de los compuestos deverbales, como en la sección anterior, presentaremos

análisis más afines a la corriente generativa y otros más compatibles con los acercamientos basados en el uso.

En la corriente generativista, la propuesta de Lardiere y Schwartz (1997), que se basa en Di Sciullo (1991, 1992), aparece en varios trabajos de adquisición de L2/L3 para explicar la formación de los compuestos en español (Desrochers et al., 2003; González Alonso et al., 2016; Liceras et al., 2004; Pomerleau, 2000; Salomaa-Robertson, 2000), en inglés (García Mayo, 2006; González Alonso et al., 2016; Lardiere, 1995; Lardiere & Schwartz, 1997; Murphy, 2000), en euskera (García Mayo, 2006; González Alonso et al., 2016) y en francés (Murphy, 2000; Pomerleau, 2000; Salomaa-Robertson, 2000), entre otras lenguas. Según esta propuesta, que aparece ilustrada en (3.120), los compuestos son cláusulas sub-sintácticas constreñidas por las mismas reglas que la sintaxis. Los indicios negativos de los nodos indican, siguiendo la notación de Cinque (1993), que se trata de una derivación sub-sintáctica y que, por lo tanto, los elementos internos del compuesto no son accesibles para la sintaxis externa.

(3.120) Estructura de los deverbales según Lardiere y Schwartz (1997)



En estas estructuras, las derivaciones de los compuestos deverbales del inglés y del español son similares. En ambos casos, se considera que el objeto tiene una forma no referencial genérica, que se realiza con la forma del plural en el *Spell-out*. Sin embargo, en inglés, la incorporación del objeto (*bottles*) al verbo (*open*) impide la realización de la forma plural del generico en el *Spell-*

out. En cuanto a la posición del verbo, como se aplican las mismas reglas que en la sintaxis de cada lengua, en inglés, el verbo no tiene que ascender como en español para cotejar su rasgo de concordancia. Consideramos que la propuesta del español se puede aplicar al francés (*ouvre-bouteille(s)*, “abre-botella(s)”) y que la propuesta del inglés se puede aplicar al euskera (*botila-irekigailu*, “botella abridor”). En euskera, como mencionan González Alonso et al. (2016), parece razonable asumir que el plural genérico del objeto queda suprimido, como en inglés, a la hora de incorporar el objeto en el verbo. En cuanto a la -s générica de los deverbales en francés, que muestra variabilidad, consideramos que se trata de una variación que ocurre en el *Spell-out* y que podría estar relacionada con la opacidad de la ortografía o con la ausencia del partitivo que normalmente aparecería en una estructura sintáctica equivalente en francés (“*ouvre des bouteilles*”).

3.3.2.1 Consideraciones semánticas

Si bien el análisis de Lardiere y Schwartz (1997) proporciona explicaciones sobre la forma de los compuestos en las tres lenguas de nuestro estudio, no permite indagar en los matices semánticos que pueden aparecer entre las diferentes lenguas. Por ejemplo, tal análisis formal no ofrece explicaciones sobre la mayor productividad de los deverbales instrumentales en comparación con los agentivos en francés y en español. Tampoco permite explicar por qué ciertos verbos parecen ser más productivos que otros para acuñar nuevos compuestos.

Desde una perspectiva basada en el uso, como los compuestos N+N, los deverbales se representan mediante diferentes construcciones que se pueden articular de maneras ligeramente distintas según el análisis elegido, como la Gramática cognitiva de construcción (Goldberg, 2006, p. 214) o la Morfología Construccional (Booij, 2005, 2009, 2010, 2013). En este marco teórico, para explicar las diferencias entre una estructura como “*botila-irekigailu*” [N+V+*gailu*] en euskera y “abrebotellas” [V+N] en español, es suficiente postular la existencia de un esquema diferente para cada lengua. Entonces, el foco de interés no se organiza entorno a la forma como tal, sino a las relaciones semánticas entre los distintos elementos de las construcciones.

En línea con los principios de la Gramática cognitiva de construcción de Goldberg (2006), Yoon (2009) propone que aun los compuestos deverbales que comparten una forma similar pueden corresponder a construcciones distintas. Por ejemplo, los compuestos deverbales de la forma V+N en español, según su significado respectivo, pueden pertenecer al tipo 1, como en (3.121), o al tipo 2, como en (3.122).

(3.121) limpiabotas, lavaplatos, quitamanchas, rompeolas (Yoon, 2009, p. 513)

(3.122) saltamontes, tientaparedes, trotacalles

Si bien los dos tipos de compuestos deverbales comparten la misma forma, se diferencian cuando se observa detalladamente el significado de cada uno. Los del tipo 1, como “limpiabotas”, evocan una escena en la que un agente ejerce una acción sobre un objeto ajeno y esa acción beneficia a un agente ajeno (p. ej., una persona X limpia las botas para un beneficiario Y; el instrumento X lava los platos para un beneficiario Y, etc.). Además, la acción representada en este tipo de compuesto tiene una duración determinada, con un principio y un fin.¹⁴⁹ En cambio, los compuestos del tipo 2 como “saltamontes” evocan un agente que hace una actividad para sí mismo, en este caso, un insecto que, metafóricamente, salta por encima de los montes, sin que el resultado de la actividad afecte a los montes. Además, esta escena tiene una duración indeterminada, frente a la acción de los deverbales de tipo 1 que, si bien es habitual, tiene un principio y un fin determinados. Entonces, para diferenciar las dos construcciones, Yoon considera que la primera tiene los rasgos semánticos [+beneficiario, +acción habitual, +nominal], mientras que la segunda tiene los rasgos [-beneficiario, +actividad habitual, +nominal]. Estas dos construcciones están a su vez relacionadas con una construcción más abstracta que tiene la misma forma [V+N], pero que tiene solamente los rasgos comunes a todos los compuestos [V+N] del español, es decir [+habitual, +nominal] (Yoon, 2009, p. 523). Las construcciones postuladas por Yoon representan solo un ejemplo de cómo se pueden representar los compuestos deverbales del español mediante construcciones, pero ilustran la importancia de considerar las diferencias semánticas además de las diferencias formales.

En cuanto a las demás lenguas de nuestro estudio, se podría postular que en francés existen, como mínimo, las mismas construcciones que en español y que en euskera existen construcciones que tienen formas diferentes, pero que cumplen papeles semánticos similares. Para explicar la variedad de compuestos deverbales en francés, Desmets y Villoing (2010) presentan un análisis

¹⁴⁹ Yoon alude aquí al contraste entre acción y actividad en el sentido de Van Valin y LaPolla (1997) que clasifican la acción como un “*State-of-affairs*” y la actividad como un “*Aktionsart*”. Referimos el lector a la obra de Van Valin y LaPolla para mayores detalles.

distinto del de Yoon (2009), con más construcciones.¹⁵⁰ Si bien la construcción más abstracta de Desmets y Villoing (2010) coincide a grandes rasgos con la de Yoon (2009), de ella se desprenden diez construcciones diferentes, todas ellas relacionadas entre sí de manera jerárquica (p. 19). De esas construcciones, las autoras afirman que la más prototípica del francés actual es la que permite crear compuestos cuyo referente semántico es un proto-agente (un agente [+animado] o un instrumento [-animado]) y cuyo N cumple el papel de paciente. Esta construcción corresponde, por ejemplo, a compuestos como “*ouvre-boite*” (“abrelatas”, instrumento) o “*gratte-papier*” (lit. “rasca-papel”, agente, manera peyorativa de referirse a un funcionario). En cambio, los compuestos con estructuras menos productivas como “*remue-méninges*” (lit. “remueve-meninges”, evento, lluvia de ideas), “*pisse-chien*” (lit. “mea-perro”, causativo, planta que hace orinar a los perros) o “*gobe-mouton*” (“traga-oveja”, planta o píldora que mata al ganado que se la traga) corresponden a otras construcciones. En euskera, si bien no parecen existir análisis de los compuestos deverbales mediante una gramática de construcciones, podríamos postular que los compuestos deverbales de esa lengua presentan un esquema similar a los del inglés. Booij (2009) considera que los compuestos deverbales del inglés corresponden a la confluencia de dos construcciones [N+N] y [V+er] que resultan en la construcción [N+V+er] (pp. 212-214). Por extensión, podemos considerar que, en euskera, los distintos esquemas formadores de deverbales ([V+*tzaile*], [V+*tzeko*], [V+*gailu*]) pueden fusionarse con el esquema [N+N] para formar un esquema de compuesto verbal como el postulado en inglés por Booij (2009). Siguiendo la lógica de las

¹⁵⁰ El análisis de Desmets y Villoing (2010) no es solamente distinto en cuanto al número de esquemas postulados, sino que se enmarca un marco teórico distinto; las autoras aplican los principios de la morfología construccional (Corbin, 1992; Fradin, 2005; Villoing, 2003, 2009) y formalizan sus esquemas mediante la gramática sintagmática guiada por los núcleos (HPSG, del inglés *Head-driven Phrase Structure Grammar*). Si bien la formalización de las construcciones en la HPSG y en la Gramática cognitiva de construcción de Goldberg (2006) son distintas, el paralelo es de interés para nuestra discusión. En ambas gramáticas, de una construcción (más abstracta) que tiene una forma dada se pueden desprender otras construcciones más específicas, como es el caso con los compuestos deverbales, que tienen la forma V+N pero que pueden corresponder a diferentes construcciones específicas en función de sus restricciones semánticas. Los detalles de la HPSG quedan fuera del marco de nuestra tesis, pero referimos al lector a Desmets, Amond y Lavieu (2003) para una introducción a esta teoría.

gramáticas de construcción, además de tener una forma distinta, cada construcción resultante tendría su significado propio, como ilustramos en (3.123).¹⁵¹

- (3.123) a. [N+V+*tzaile*, un agente V+*tzaile* realiza una acción V en relación con N]
kale- garbitzaile
calle- limpiador
“limpiacalles”
- b. [N+V+*gailu*, un instrumento V+*gailu* realiza una acción V en relación con N]
lursagar- zurigailu
patata- pelador
“pelapatatas”
- c. [N+V+*tzeko*, un instrumento V+*tzeko* realiza una acción V en relación con N]
lata- irekitzeko
lata- abridor
“abrelatas”

Volviendo de manera más general a los principios de los acercamientos basados en el uso y las construcciones postuladas para los compuestos deverbales del español, recordemos que las construcciones más frecuentes y prototípicas tienen más probabilidades de servir para instanciar nuevos ejemplares, que es lo que observamos en el uso. Si bien los dos esquemas presentados por Yoon están disponibles en español, el que tiene los rasgos [+beneficiario, +acción habitual, +nominal] es el más frecuente y, por lo tanto, el que genera la mayor cantidad de nuevas construcciones. Así, según los datos de Lacroix (en preparación), 64 compuestos deverbales incorporados al DLE en 2014 correspondían a la construcción de tipo 1, frente a solo 11 construcciones para el tipo 2. El efecto de frecuencia no solo tiene un impacto en la construcción abstracta, sino también en los verbos utilizados para formar los compuestos. Así, 31 de esas 75 nuevas incorporaciones al DLE están formadas a partir de los 15 verbos más frecuentes en los compuestos deverbales del español recopilados en todo el DLE. Se observa un efecto comparable con verbos más frecuentes en los compuestos del francés, como “*porter*”, “*couper*”, “*garder*” (“portar”, “cortar”, “guardar”) y del euskera, como “*eman*”, “*jan*” (“dar”, “comer”), en los datos recopilados por Lacroix (en preparación).

¹⁵¹ Como hemos mencionado antes, dos construcciones pueden tener un mismo significado, pero presentar una forma distinta, como ocurre entre (3.123b-c). También es posible que existan matices semánticos que no distinguimos dada la profundidad del nivel de análisis que adoptamos ahora.

En resumen, en la presente sección, hemos presentado distintos análisis de los compuestos deverbales, algunos de los cuales otorgan más importancia al carácter sintáctico de los compuestos, mientras que otros se basan en las construcciones y consideran tanto el significado como la forma a la hora de definir las estructuras de los compuestos. Los primeros se adscriben a la corriente generativista, mientras que los segundos son más afines a los enfoques basados en el uso. Ambas posturas permiten explicar las características formales de los compuestos (direccionalidad, flexión verbal o derivación, -s générica), mientras que las propuestas basadas en las construcciones permiten añadir matices semánticos adicionales. En el marco de nuestro estudio, los matices aportados por las construcciones y los principios de analogía y de prototipicidad más propios de los enfoques basados en el uso podrían ser útiles para discernir la influencia de una lengua frente a otra. A este respecto, volveremos sobre las hipótesis precisas de nuestro estudio tras la revisión de literatura de los estudios de adquisición de los compuestos en L1/2L1/L2/L3, la cual presentamos a continuación.

3.4 Adquisición de los mecanismos de composición

En la presente sección, repasaremos los principales trabajos que se han llevado a cabo en las distintas ramas del campo de la adquisición de lenguas. Los trabajos de adquisición de L1 nos servirán de base para contrastar las dificultades que encuentran los hablantes en su L1 con las encontradas en adquisición de 2L1 o de L2. Finalmente, terminaremos con los estudios sobre adquisición de L3 para ilustrar, entre otras cosas, cómo las interacciones entre las lenguas se vuelven más complejas cuando incluimos una tercera lengua en el repertorio lingüístico del aprendiente.

3.4.1 En L1

Los niños tienden a usar el proceso de composición más temprano en lenguas como el inglés que en otras en las cuales se trata de un proceso menos productivo y, dentro del inglés, los compuestos más frecuentes (N+N) se adquieren antes que los menos frecuentes, como los deverbales (Clark, 1998). Clark y Hecht (1982) apuntan, además, que el valor semántico del referente también parece jugar un papel, ya que los niños anglohablantes tienden a usar los compuestos deverbales para referirse a agentes [+animado] antes de generalizar ese uso a los instrumentos [-animado]. Por su parte, la complejidad estructural de los compuestos también parece

influir en el tiempo de adquisición, es decir, la simple yuxtaposición de dos sustantivos (compuestos N+N) se adquiere antes que otros procesos que requieren el uso de mecanismos más complejos (flexión verbal, colocación de un interfijo, derivación, etc.). A modo de ejemplo, se considera que los niños de inglés L1 empiezan a combinar dos raíces para formar nuevas palabras (los llamados *root compounds*) a partir de los 2 años de edad (Clark et al., 1985), pero su capacidad para producir compuestos deverbales (N+V+er) solo aparece hacia los cinco o seis años de edad (Clark et al., 1986; Nagpal & Nicoladis, 2009). Además, los niños hablantes de inglés L1 no suelen tener problemas de direccionalidad en los compuestos N+N (Nicoladis, 1999), pero en cuanto a los N+V+er, pasan por una etapa de desarrollo en la que invierten el orden de los componentes (p. ej., *thrower ball*, en lugar de *ball thrower*; Barbosa & Nicoladis, 2016; Clark et al., 1986; Nagpal & Nicoladis, 2009). La morfología de plural es otro tema que ha recibido mucha atención en la adquisición de los compuestos N+V+er del inglés. Si bien los compuestos deverbales del inglés no suelen incluir una -s genérica en el complemento del verbo (p. ej. **balls-thrower*) como se hace en español (p. ej., lanzapelotas), existe la posibilidad de formar un compuesto con el complemento en plural en casos de plurales irregulares (p. ej., *mice eater* o *mouse eater* son ambos aceptables). En un experimento con niños de anglohablantes de entre 3 y 5 años de edad, Gordon (1985) observó que éstos no producían un plural interno con los sustantivos de plural regular en un 98 % de los casos, pero que en un 98 % de las respuestas que contenían un sustantivo de plural irregular, la forma plural sí aparecía.¹⁵²

En adquisición del español como L1, en un estudio con niños hispanohablantes de 3, 4 y 5 años, Álvarez et al. (1999) observan que, si bien los niños usan los compuestos deverbales que ya tienen lexicalizados a partir de los tres años, solo empiezan a usar la composición para denominar

¹⁵² Si bien, para los adultos, la forma del plural no es la más frecuente en los compuestos N+V+er que contienen sustantivos con plural irregular, se trata de una opción posible. Según Gordon (1985), esta posibilidad procede de lo que llama la teoría de las reglas ordenadas por niveles (*Level-ordered rules*) en el léxico. Diferentes operaciones léxicas ocurren en diferentes niveles, es decir: 1) las derivaciones irregulares e idiosincrásicas que alteran la palabra y que no son productivas (como los plurales irregulares), 2) derivaciones regulares que no alteran la palabra y que son productivas y más predecibles (como la formación de compuestos) y 3) flexiones regulares. Así, es posible crear un compuesto a partir de “*mouse*” o “*mice*” (nivel 1), pero no “*mouses*” (nivel 3), ya que la formación del compuesto es de nivel 2. Según estas reglas, después de la formación de un compuesto, solo es posible la flexión plural al final de este (*mice-eaters*, *mouse-eaters*) y esta marca de plural tiene un abarque sobre todo el compuesto como conjunto, no sobre “*mouse*”.

instrumentos inventados a partir de los 4 años y, de manera más productiva, a los 5 años. En cuanto a la -s genérica que suele aparecer en los compuestos deverbales del español, los autores no presentan cifras precisas, ya que no se trataba del foco del estudio, pero en los ejemplos que reportan podemos notar una mayoría de ejemplos con -s (p. ej., “abrebotellas”, “sacapuntas”, “pintalabios”, “matamoscas”, “cortapapas”, “sacapapas”, frente a “pelazanahoria” y “sacapintura”).

Existen pocos datos sobre las etapas de adquisición de los compuestos N+N en francés como L1, pero Clark (1998) afirma que los niños francohablantes de 5-6 años de edad suelen preferir otros mecanismos para modificar un sustantivo, como los sintagmas preposicionales. Nicoladis (1999) estudió las producciones de un niño entre 2;9 y 3;3 años de edad a partir de los datos de Suppes et al. (1973) y descubrió que el niño producía compuestos ya presentes en su *input*, pero también compuestos noveles, sobre todo a partir de los tres años de edad. Así, a pesar de que no se considere la composición N+N en francés como un mecanismo productivo, los niños francohablantes parecen recurrir a ello un año más tarde que los que hablan inglés. En cuanto a los compuestos deverbales, Clark (1998) observa que el morfema derivativo *-eur* que se puede usar en sintagmas como “*vendeur de livres*” (“vendedor de libros”, con significado similar a “vendelibros”) se adquiere hacia los tres años y medio de edad, y basándose en los datos de Seidler (1988), afirma que la composición deverbal es un mecanismo poco usado para acuñar nuevos instrumentos o agentes. En cambio, Nicoladis (2005) observa que los compuestos deverbales son productivos en francés a partir de los 4-5 años.

En cuanto al euskera, si bien existen obras descriptivas del fenómeno como hemos visto en la sección 3.2.3, no hemos encontrado ningún estudio de adquisición del fenómeno en L1.¹⁵³ Sin embargo, basándonos en el inglés, que muestra características semejantes en este tipo de estructuras, podemos suponer que la formación de compuestos N+N aparece antes de los

¹⁵³ La mayoría de los trabajos de adquisición que incluyen el euskera no son de adquisición de L1 como tal, sino de adquisición bilingüe (simultáneos o secuenciales tempranos; francés/euskera o español/euskera), por cuestiones de la realidad sociolingüística del País Vasco. Por otra parte, si bien los trabajos de adquisición (bilingüe) son numerosos, suelen centrarse en las particularidades morfosintácticas del euskera, como la ergatividad (Duguine, 2013; Duguine et al., 2014; M.-J. Ezeizabarrena, 2012) o las marcas casuales de manera más general (Ezeizabarrena Segurola et al., 2009), los sujetos y los objetos (Azpiroz & Ezeizabarrena Segurola, 2011, 2012; M. J. Ezeizabarrena, 2013) y la morfología verbal (Ezeizabarrena Segurola, 1996, 1996), entre otros temas.

compuestos deverbales y que ambos mecanismos de composición ya han sido adquiridos hacia los 4-6 años de edad.

De manera general, la complejidad estructural juega un papel en los primeros años de desarrollo de los mecanismos de composición y hace variar ligeramente la edad adquisición de cada estructura dentro de una misma lengua, pero también de una lengua a otra. Sin embargo, después de esas primeras etapas de adquisición del lenguaje, a partir de los 5-6 años, las dificultades estructurales desaparecen y, luego, la frecuencia de uso y la tipología de la lengua meta cobran más importancia (Berman, 2009).

3.4.2 En 2L1

En adquisición bilingüe, Nicoladis (1999) se interesó por la direccionalidad de los compuestos N+N en la producción natural de un niño bilingüe francés-inglés entre los 2;9 y 3;3 años de edad. Sus resultados sugieren que, a pesar de los errores de direccionalidad cometidos en sus dos lenguas, el niño es capaz de diferenciar las reglas de cada lengua, ya que muestra una mayoría de compuestos N+N en francés con el núcleo inicial, y N+N en inglés con el núcleo final. Un estudio posterior de Nicoladis (2002c) con tareas experimentales puso de relieve las diferencias entre la comprensión y la producción de los compuestos N+N. Al comparar niños bilingües francés-inglés y monolingües inglés de 3-4 años, Nicoladis observó más inversiones de direccionalidad en la tarea de producción que en la tarea de comprensión. De hecho, con respecto a la comprensión, no observó ninguna diferencia significativa entre el grupo monolingües y los grupos bilingües.

Por su parte, Garza-González (2014) se interesó por la interpretación de los compuestos V+N y N+N en hablantes de español como lengua de herencia en los Estados Unidos (bilingües adultos inglés-español).¹⁵⁴ Los participantes pertenecían a 3 grupos experimentales: bilingües tardíos (inglés L2, después de los 12 años), bilingües tempranos (inglés L2, entre los 3 y 6 años), y bilingües simultáneos (con el inglés y el español adquiridos antes de los 3 años). También un grupo de control de hablantes de español L1 (de México) participó en el estudio. El experimento, un cuestionario de opción múltiple, tenía por objetivo evaluar los errores de interpretación en la

¹⁵⁴ Por cuestiones de concisión, solo comentamos los resultados del experimento 1, que tenía que ver con la direccionalidad, ya que es de interés particular para nuestro estudio. El experimento 2 se relacionaba con la transparencia de los compuestos, un concepto que no abordamos en nuestra tesis.

direccionalidad de compuestos noveles. Por ejemplo, para interpretar el compuesto "campo turista" tenían que elegir entre a) campo para turistas o b) turista en el campo. En este caso, la opción (b) sugería una interpretación errónea con un núcleo a la derecha como en inglés, mientras que la opción (a) sugería una interpretación correcta, con un núcleo inicial.¹⁵⁵ Cabe mencionar de entrada que los promedios de acierto son bastante altos para todos los grupos experimentales, incluso el de bilingües simultáneos, que obtuvo los resultados más bajos (80 %). Los bilingües simultáneos cometieron más errores de interpretación, seguidos por los bilingües tempranos y los tardíos, respectivamente. Los bilingües tardíos no se diferenciaron de manera significativa del grupo de control de hablantes nativos de español y todos los grupos experimentales cometieron menos errores de interpretación con los V+N que los N+N. La autora concluye que la frecuencia parece tener un papel importante en los errores de los participantes. Por una parte, los V+N son más frecuentes que los N+N en español y, por consiguiente, todos los participantes han recibido más *input* de los primeros. Por otra parte, los bilingües que han adquirido el inglés más tarde (simultáneo < temprano < tardío) tienen menos dificultades con los compuestos del español, ya que han recibido más *input* en español. Además de la frecuencia, con respecto a los mejores resultados con los V+N que con a los N+N, la autora añade que el orden Verbo+Objeto presente en los compuestos deverbales de las lenguas romances también se encuentra en una etapa de adquisición de los compuestos deverbales del inglés (la etapa en la que los niños producen compuestos invertidos como *thrower ball* que comentamos antes) y que podría haber tenido un efecto facilitador en el proceso de adquisición de la estructura en español. Queremos proponer como alternativa que la prominencia de la forma también podría haber tenido un peso importante. Concretamente, para los N+N, la forma superficial es la misma en inglés y en español, un sustantivo seguido por otro; no hay ningún indicio en la forma que permita identificar el núcleo y, por lo tanto, establecer una direccionalidad clara. En cambio, para los compuestos deverbales (V+N/N+V+*er*), el hecho de que los dos elementos tienen una forma distinta (un verbo conjugado o derivado, frente a un sustantivo) puede contribuir a la identificación del núcleo. Encontramos un indicio de esto en las respuestas obtenidas para ciertos ítems, en los que el verbo conjugado en tercera persona es homónimo de un sustantivo o de un adjetivo. Los dos ítems que provocaron más errores de interpretación son "aguafiestas" (en el que el sustantivo "agua" es más frecuente que la conjugación del verbo

¹⁵⁵ De manera similar, los ítems [V+N] consistían en compuestos como "montacargas", con dos opciones de interpretación: (a) vehículo para montar cargas en plataforma o (b) cargas que se montan en una plataforma.

“aguar”) y “limpiabotas” (en el que el adjetivo “limpio/limpia” es bastante frecuente también), por lo que es posible que ciertos participantes no hayan podido identificar que se trataba de un compuesto V+N e interpretar de la manera esperada el compuesto.¹⁵⁶

En resumen, en adquisición bilingüe, si bien la interacción entre las dos lenguas puede provocar tasas de errores más altas en las primeras etapas de adquisición, los niños bilingües son capaces de diferenciar las características de los compuestos en sus dos lenguas. Como en la adquisición de L1, después de las primeras etapas de adquisición, la frecuencia juega un papel importante en el nivel de dominio de los mecanismos de composición.

3.4.3 En L2

Como se ha mencionado antes, los compuestos N+N y deverbales son los compuestos más estudiados en adquisición de L2. En esta sección, revisaremos trabajos que se enfocan exclusivamente en los N+N, como Liceras y Valenzuela (1998), Liceras y Díaz (2000b), Liceras et al. (2002), Slabakova (2002), Trías y Villanueva (2011, 2013), De Cat et al. (2015) y Liceras y Klassen (2019), o exclusivamente en los deverbales, como Lardiere y Schwartz (1997), Murphy (2000) y Desrochers et al. (2003). Incluiremos también el trabajo de Salomaa-Robertson (2000) que, como es el caso de la presente tesis, cubre tanto los compuestos N+N como los deverbales. Empezaremos con la revisión de los estudios sobre los compuestos N+N y continuaremos con los datos sobre los deverbales.

En adquisición del español como L2, Liceras y Valenzuela (1998) estudiaron la adquisición del parámetro de la composición (Snyder, 1995) por parte de hablantes de francés L1 [-TCP] o inglés L1 [+TCP] que aprendían el español como L2 [-TCP].¹⁵⁷ Los resultados de una prueba de juicios de gramaticalidad revelaron que, en los niveles inicial e intermedio de español, los

¹⁵⁶ En el *Corpus del español: NOW* (Davies, 2019), el sustantivo “agua” (en singular) tiene una frecuencia relativa de 294.78 (por millón de palabras) mientras que el verbo “aguar” conjugado en tercera persona ni siquiera aparece una vez (o si es el caso, no ha sido correctamente identificado). Por su parte, el adjetivo “limpia” (en femenino singular) tiene una frecuencia de 11.49 frente a 1,06 para el verbo limpiar conjugado en tercera persona.

¹⁵⁷ Si bien en la sección 3.1.7 argumentamos que la diferencia entre estas lenguas no era tan categórica, utilizamos la clasificación binaria de Snyder (Snyder, 1995) en función del parámetro de la composición (\pm TCP, *The Compounding Parameter*) porque representa el punto de referencia de los autores del estudio. Esto se aplica también a la mayoría de los estudios que revisaremos en la presente sección.

francófonos eran capaces de rechazar con más acierto que los anglófonos los compuestos agramaticales, lo que parecía indicar un efecto de la L1 en la L2. Las autoras también notaron que este efecto desaparecía en los niveles avanzados, ya que ambos grupos podían rechazar con gran acierto las construcciones agramaticales. Liceras y Díaz (2000b), en un estudio sobre los desencadenantes de la adquisición de los compuestos N+N en español L2 por parte de hablantes de diferentes L1 (francés, inglés, alemán, ruso, polaco, danés, sueco, chino, japonés y coreano), revelaron que los aprendientes de español L2 no parecían tener acceso al marcador de palabra (J. W. Harris, 1991; Piera, 1995), por lo menos en los niveles inicial e intermedio, pero que sí parecían adquirir la direccionalidad correcta (núcleo inicial), al menos en los niveles intermedio y avanzado. Para llegar a esa conclusión sobre el marcador de palabra, las autoras se apoyaron en las marcas erróneas de género encontradas en los compuestos (“*araño hombre” o “mujer *pulpa” por ejemplo). Recordemos que la noción de marcador de palabra está originalmente vinculada con el género (J. W. Harris, 1991) y que, en los compuestos N+N en español, solo el núcleo lleva marcas de género y número.¹⁵⁸ Entonces, los ejemplos agramaticales encontrados por Liceras y Díaz (2000) muestran un problema de acceso al marcador de palabra en español L2. En cuanto a la direccionalidad, las autoras atribuyen el progreso de los aprendientes en los compuestos a un mejor dominio de la direccionalidad de su español L2 de manera general.

Slabakova (2002) también estudió el parámetro de la composición en español L2 con hablantes de dos L1 diferentes, francés [-TCP] e inglés [+TCP]. En lo que concierne a la direccionalidad de los compuestos N+N, los resultados de Slabakova van en el sentido de los resultados anteriores reportados en los estudios anteriores de Liceras y colegas, es decir, una influencia de la L1, positiva en el caso del francés y negativa en el caso del inglés. El efecto de esta influencia, como en los estudios de Liceras y colegas, disminuye a medida que los aprendientes avanzan en los niveles intermedios y avanzados. En lo que concierne a las demás propiedades vinculadas con el parámetro de la composición de Snyder, los resultados no apuntan hacia un efecto de conjunto (*cluster*) de las diferentes propiedades, ya que el dominio de una propiedad no parece desencadenar el dominio de las demás propiedades del macro-parámetro.

Trías y Villanueva (2011) también se interesaron por la configuración del parámetro de la composición por parte de estudiantes de L2 español en dos contextos de adquisición distintos: uno

¹⁵⁸ Más información sobre el marcador de palabra (MP) en la sección 3.3.1.2 de este capítulo.

en el que los participantes tenían clases de español, pero además vivían en inmersión (Caracas, Venezuela) y otro en el que los aprendientes solo tenían clases de español en el que la lengua fuera del aula era el inglés (Michigan, EE.UU). Además de la diferencia de contextos, los participantes tenían distintas L1, que las autoras usaron para clasificar a los participantes según la configuración del parámetro de la composición de Snyder (1995, 2001). En una tarea de elicitación a partir de una paráfrasis y una imagen, las autoras evaluaron el tipo (N+N, N+de+N) y la direccionalidad de las respuestas proporcionadas. No se observó una diferencia significativa entre los grupos según su L1 para la producción errónea de compuestos NN con el núcleo a la derecha. Considerando que todos los participantes eran principiantes, la alta tasa de producción de compuestos N+N con la direccionalidad correcta contrasta con la influencia de la L1 observada en los estudios anteriores que hemos presentado. Sin embargo, queremos subrayar que la alta tasa de respuestas con la direccionalidad correcta se podría atribuir simplemente a un efecto de la tarea. Si bien solo aparece un ejemplo de ítem en Trías y Villanueva (2011), nos parece que si todos los demás ítems están contruidos siguiendo la misma forma, éstos podrían fomentar la producción de compuestos con un núcleo a la izquierda, simplemente por la formulación de la paráfrasis: “Este es un sofá (imagen de sofá) y este es un ... (imagen de sofá cama)”. A pesar de ello, las autoras encontraron una diferencia significativa en función del contexto de aprendizaje: el grupo de Venezuela produjo una tasa significativamente más baja de compuestos con la direccionalidad inversa. En cuanto a las respuestas con la direccionalidad correcta, las autoras observaron que los participantes de que tenían una L1 [+TCP] producían significativamente más respuestas del tipo N+de+N que los hablantes de L1 [-TCP]. Para las autoras, esta diferencia se podría atribuir al hecho de que los participantes han categorizado el español como una lengua en la que los compuestos NN constituyen una opción menos productiva y, por lo tanto, optan por la estructura con preposición.

En un estudio posterior con los mismos grupos de participantes, Trías y Villanueva (2013) presentan los resultados de una prueba de juicio de gramaticalidad adaptada de Slabakova (2002) y una de opción múltiple en la cual los participantes tenían que elegir entre respuestas [N+N] y [N+de+N] con la direccionalidad correcta o incorrecta. Por ejemplo, para la oración “Es importante salir antes de las 6:00 de la tarde para evitarnos el tráfico de la ...”, podían escoger entre (a) pico hora, (b) hora pico o (c) hora de pico. En la prueba de juicios de gramaticalidad, las autoras observaron un efecto significativo de grupo según la L1: los participantes con una L1 [-TCP] (como el francés) obtuvieron una tasa significativamente más alta de aciertos. En cuanto a la prueba de

opción múltiple, los participantes del grupo L1 [+TCP] (como el inglés) escogieron significativamente más respuestas de tipo [N+N] con el núcleo a la derecha que los del grupo L1 [-TCP]. Además, el contexto de aprendizaje parece interactuar con la L1 de los participantes ya que los que tienen una L1[+TCP] y que aprenden el español en EE. UU. son los que contestan con más respuestas de núcleo a la derecha. En cuanto a la diferencia entre la producción de [N+N] y [N+de+N] con un núcleo a la izquierda, los resultados son difíciles de interpretar, ya que la respuesta esperada era del tipo [N+de+N] para 13 de los 17 ítems y del tipo [N+N] para los cuatro ítems restantes.

De estos dos estudios queremos destacar dos aspectos, la L1 parece influir en la direccionalidad de los compuestos en español L2 y el *input* parece modular esta influencia. Concretamente, los participantes tienden a reproducir la direccionalidad de su L1 en los compuestos en español L2. Si bien esta diferencia es poco perceptible en algunos casos, hemos visto que esta ausencia de efecto se podría atribuir a un efecto de las tareas experimentales. En cuanto al *input*, los resultados muestran un efecto positivo del contexto de aprendizaje en Venezuela donde los participantes no solo tienen instrucción formal, sino que viven en inmersión y, por lo tanto, tienen acceso a un *input* más diversificado que contiene potencialmente más compuestos con el núcleo inicial. Además, como viven en inmersión, tienen un refuerzo constante del *input* sobre el hecho de que, en español, el núcleo del sintagma nominal suele aparecer a la izquierda. Sin embargo, es posible también que este contexto tenga un efecto positivo más global sobre el nivel de dominio del español y que, en realidad, los participantes del grupo de Venezuela no sean de nivel comparable a los de los EE. UU. Sin embargo, la falta de información sobre el nivel de español de cada grupo no permite corroborar esta hipótesis.

De Cat et al. (2015) investigaron las diferencias de procesamiento entre hablantes nativos de inglés y hablantes de nivel muy avanzado de inglés L2 que tenían como L1 una lengua de núcleo inicial (español) o núcleo final (alemán). Los autores administraron dos tareas de decisión léxica en las cuales los participantes tenían que aceptar o rechazar compuestos correctamente formados (p. ej., “coal dust”) o con el orden inverso (p. ej., “dust coal”). La primera tarea estaba cronometrada y se medía en tiempo de reacción, en la segunda, los participantes tenían un tiempo fijo para cada ítem y las respuestas cerebrales (potencial relacionado con eventos o *Event-related Potential*; ERP) se medían mediante electroencefalogramas (EEG). En ambas tareas se medían también los porcentajes de acierto en la aceptación o rechazo de los compuestos y no compuestos. Los autores

observaron que los participantes de L1 español mostraban una tendencia a aceptar más los compuestos con el orden inverso, pero esta diferencia entre los grupos no era significativa. Dado el nivel avanzado de inglés de todos los participantes, estos resultados no son sorprendentes, todos parecen haber adquirido los mecanismos de composición de la lengua meta. Sin embargo, se observó un tiempo de reacción mayor con los participantes de L1 español, lo cual es interpretado por los autores como una influencia de la L1, dado que los participantes de alemán L1 no muestran el mismo efecto. Entonces, la influencia de la L1 no afecta de manera significativa la capacidad de aceptar o rechazar los compuestos de la lengua meta, pero cuando se observa más en detalle, con la velocidad de procesamiento, la influencia de la L1 parece perdurar hasta el nivel avanzado de los hablantes de L2 de nivel avanzado. Los resultados del ERP van en el mismo sentido: los participantes de español L1 mostraron un efecto de frecuencia de los compuestos con los ítems que correspondían al orden de su L1 (es decir, compuestos frecuentes en inglés, pero con el orden inverso), el efecto contrario a lo que se observó con el grupo de control de inglés L1, lo cual es interpretado por los autores como una confirmación de la influencia de la L1.

Más recientemente, Liceras y Klassen (2019) se interesaron también por los compuestos en español L2, pero más concretamente por la interacción entre composición y derivación e integraron el uso de diminutivos en compuestos N+N en español (p. ej., “hombrecito rana”). Las autoras observaron diferencias en el abarque que los hablantes atribuían al diminutivo. Mientras que los nativos preferían la adjunción del diminutivo al núcleo del compuesto y le atribuían una interpretación que abarcaba a todo el compuesto, los no nativos (inglés L1) tenían tendencia a restringir el abarque del diminutivo al elemento al que estaba adjunto, ya fuera el núcleo o el modificador. Por otro lado, las autoras observaron también un efecto del rasgo [\pm animado] en sus resultados. Cuando se trataba de un compuesto de referente inanimado (p. ej., “carta bomba”), la percepción de separabilidad de los dos componentes era mayor que cuando se trataba de un referente animado (ej. “hombre lobo”). Concretamente, el abarque del diminutivo parecía ser mayor en los compuestos animados que en los compuestos inanimados. Por ejemplo, se suele interpretar el compuesto “hombrecito lobo” como un ser híbrido pequeño más que un ser con un cuerpo de hombre pequeño y una cabeza de lobo de tamaño normal. En cambio, como la percepción de separabilidad es mayor en los compuestos inanimados, las dos interpretaciones del abarque del diminutivo son posibles para una “cartecita bomba”.

En lo relativo a los compuestos deverbales, Lardiere y Schwartz (1997) estudiaron los compuestos N+V+er en inglés como L2 por parte de hablantes de español como L1. Las autoras presentan evidencias de influencia a partir de la L1, es decir, basándose en la direccionalidad y en la morfología de su L1, los participantes producen compuestos idiosincrásicos como “*painter-toes*” en lugar del correcto “*toe-painter*”. Si bien el efecto de la direccionalidad de la L1 es robusto en los participantes de nivel inicial (41,55 % de inversión), este tiende a decaer en los niveles intermedio y avanzado (22,65 % y 0,53 % respectivamente). En cuanto a la morfología de número, las marcas erróneas de plural persisten hasta el nivel avanzado. Incluso en los compuestos que respetan la direccionalidad del inglés (p. ej., **hands-washer*), la tasa de pluralización de los objetos en este grupo es de un 33,33 % si solo contamos los plurales regulares, pero asciende a un 52,2 % si incluimos los plurales irregulares.

Murphy (2000) también se interesó por la morfología de número en los compuestos deverbales N-V-er en inglés L2, esta vez por parte de hablantes de francés L1. A partir de preguntas como “*What do you call a person who breaks bones?*” (“¿Cómo llamas a una persona que rompe huesos?”), la autora elicitó compuestos deverbales como “*bone(*s) breaker*” (“rompe huesos”) para averiguar si los participantes usaban una marca de plural interna en los compuestos deverbales del inglés. Considerando que las marcas de plural sobre el complemento del verbo en los compuestos deverbales del francés son menos frecuentes que en español, la autora esperaba resultados diferentes de los de Lardiere y Schwartz (1997) en cuanto a la posible influencia de la L1 sobre la marca del plural. Sin embargo, la autora observó una tasa de pluralización de 46 %. Esto sugiere que, aunque la marca de plural se aplique con menos sistematicidad en francés que en español, esta sigue siendo parte integrante del proceso de composición de los deverbales y, como tal, es susceptible de ser transferida a la L2.

Desrochers et al. (2003) se interesaron por las diferencias de procesamiento de los compuestos deverbales entre hablantes nativos de español con o sin formación en filología hispánica y hablantes de español como L2 de nivel avanzado inscritos en un programa de filología hispánica (inglés o francés como L1). En una tarea de decisión léxica cronometrada que contaba con compuestos gramaticales (quitamanchas) y con compuestos agramaticales por su direccionalidad inversa (*manchasquita), por contener un verbo en infinitivo (*quitarmanchas), o por omisión de la -s final (*quitamancha), los autores midieron los tiempos de reacción y las tasas de error. De manera general, no se observó ninguna diferencia significativa entre los grupos de

nativos. En cuanto a los no nativos, si bien los autores no proporcionan ningún análisis estadístico de la diferencia entre ese grupo y los grupos nativos porque sus resultados son demasiado heterogéneos, se puede observar una tasa de error más alta y un tiempo de reacción más lento para los no nativos. Los autores reportan además una diferencia significativa del tipo de ítem sobre la tasa de errores y el tiempo de reacción: los participantes tuvieron más dificultad en rechazar los ítems sin -s (61 % de errores) que en aceptar los ítems gramaticales (25 % de errores). Además, cometieron un 12 % de errores con los compuestos de direccionalidad inversa y un 11 % de errores con los compuestos formados a partir de un infinitivo. Cabe mencionar aquí que la tarea presentó cierta dificultad para los nativos también, ya que estos produjeron un 54 % de deverbales sin -s final. Los autores apuntan que, al contrario de la forma verbal (3sg) y la direccionalidad (V+N), la -s final en los compuestos del español presenta cierta variabilidad.¹⁵⁹ Por eso, es posible que sea más difícil de procesar en una tarea que impone más presión sobre el participante, como una tarea de decisión léxica cronometrada.

Por su parte, Salomaa-Robertson (2000) estudió la adquisición de los compuestos N+N y deverbales en español L2 por parte de hablantes nativos de finés. Los resultados relativos a los N+N fueron incorporados al trabajo de Liceras et al. (2002) que presentaremos en la próxima sección, por lo que nos limitaremos aquí a comentar sobre la flexión verbal y la -s genérica en los compuestos deverbales. En Salomaa-Robertson (2000), los errores vinculados a esas dos características persisten en la producción de los aprendientes de nivel avanzado, con un 40 % de deverbales sin -s final, y una variedad de elementos idiosincrásicos que corresponden, principalmente, a: (1) formas conjugadas con una terminación errónea (-a, -e, -o), (2) verbos en infinitivo o (3) formas derivadas mediante un sufijo agentivo, como -dor o -ero.

3.4.4 En L3

Los trabajos sobre los compuestos en adquisición de L3 son más escasos, pero comentaremos las tendencias observadas en los pocos estudios encontrados. De hecho, los trabajos de Liceras et al. (2002) y Pomerleau (2000) que presentamos a continuación no se refieren al

¹⁵⁹ Si bien la -s final en los deverbales del español presenta cierta variabilidad (ver sección 3.2.1.2), recordamos al lector que es mucho más sistemática que en los deverbales en francés (sección 3.2.2.2.). Por lo tanto, en una tarea equivalente en francés, consideramos que los hablantes de francés L1 tendrían una tasa de error más elevada que lo que se observó en Desrochers et al. (2003).

español como una L3 específicamente, pero en ellos se insiste en el hecho de que varios de los participantes tienen otra L2 además del español.¹⁶⁰

En Liceras et al. (2002), las autoras presentan conclusiones similares a las observadas en los estudios anteriores de adquisición de L2. A partir de los resultados de tres grupos con distintas L1 (francés, inglés y finés), las autoras notan cierto efecto de la L1 en el dominio de la direccionalidad de los compuestos N+N en español L2. Sin embargo, la L1 no parece explicar todo, ya que el grupo de inglés L1 se comporta de manera distinta al de finés L1, a pesar de compartir la misma direccionalidad en los compuestos N+N (núcleo final) en su L1. Las autoras proponen que esta diferencia se podría atribuir al hecho de que todos los participantes del grupo de L1 inglés tienen cierto conocimiento del francés y que los del grupo de L1 francés tienen cierto dominio del inglés. En este caso, el español no es la L2 de los participantes sino su L3, lo que supone, como hemos visto en el capítulo 1 (sección 1.3), que existen otras fuentes posibles de influencia en el proceso de adquisición. Dado que el trabajo Liceras et al. (2002) no se articulaba en torno a la noción de AL3, carecemos de datos específicos sobre el conjunto de lenguas habladas por los participantes (nivel de dominio, orden de adquisición, etc.), por lo que resulta difícil profundizar en el análisis de los efectos potenciales de la L1 o la L2 en sus datos.

De manera similar, Pomerleau (2000) estudió, mediante tareas de producción y de juicios de gramaticalidad, el dominio de los compuestos N+N y deverbales del español por parte de hablantes de francés L1 (e inglés L2) aprendientes de español como L3 en niveles inicial, intermedio y avanzado. Un grupo de hablantes de español L1 también realizó las mismas pruebas y ejerció de grupo de control. En la tarea de producción, los aprendientes de todos los niveles produjeron una mayoría de compuestos N+N con núcleo inicial. Sin embargo, en comparación con el grupo de control, produjeron más formas inesperadas, como un sustantivo solo o una traducción literal del francés (p. ej., estructuras N+de+N). En los juicios de gramaticalidad, los participantes tendieron a aceptar todos los compuestos independientemente de la direccionalidad, aunque esta tendencia decae en función del nivel. Pomerleau indica que es posible que los principiantes no

¹⁶⁰ En el caso de Pomerleau (2000), todos los participantes tienen el francés como L1 y han estudiado el inglés como L2 en la primaria, la secundaria y el colegio, al ser obligatorio en la provincia de Quebec, donde se llevó a cabo el estudio.

hayan entendido bien las oraciones de la prueba de juicio de gramaticalidad y que hayan aceptado casi todo por defecto. En cuanto a los compuestos deverbales, Pomerleau observa una proporción considerable de respuestas que no corresponden a la estructura V+N sino a otras como V+P+N (“besar de pescado” en lugar del compuesto inventado “besapescados”). A pesar de ello, no observa problemas mayores con respecto a la direccionalidad, pero sí observa dificultades con la morfología verbal y de número. Los errores en la morfología verbal decaen a medida que aumenta el nivel de los aprendientes, pero los errores de morfología de número parecen más persistentes. En cualquier caso, la autora indica que la influencia del francés L1 parece ayudar a los aprendientes a establecer la direccionalidad adecuada desde el nivel inicial, al menos en lo referente a la producción. En cambio, el inglés L2, que implicaría una direccionalidad inversa, no parece influir en la producción de los aprendientes.

García Mayo (2006) se interesó por los compuestos deverbales en inglés L3 por parte de bilingües español-euskera de tres diferentes niveles de inglés. En una tarea de elicitación escrita de compuestos inventados (p. ej., “*hand-washer*”, lit. “mano-limpiador”, “una persona que limpia las manos”), la autora observó que ninguno de los participantes mostraba dificultad con la direccionalidad, pero que sí exhibían un uso divergente del plural. En concreto, ninguno de los participantes produjo un compuesto con la direccionalidad inversa que correspondería a la del español (p. ej., “**washes-hands*”), pero el 61.2 % de los participantes incluyó una -s en el elemento no-nuclear del compuesto (p. ej., “**hands-washer*”). En cuanto a la direccionalidad, estos resultados contrastan con estudios sobre inglés L2 como el de Lardiere y Schwartz (1997) que mostraban una influencia del español L1 en la direccionalidad de los compuestos producidos en inglés L2 (p. ej., “**washes-hands*”). En cuanto a la -s interna, su uso tiende a decaer en función del nivel de inglés, de manera similar a lo que se observó en Lardiere y Schwartz. García Mayo concluye que las dos lenguas previas de los participantes parecen influir en los mecanismos de composición de la L3: por una parte, el euskera contribuye de manera positiva en la direccionalidad y la derivación y, por otra, el español contribuye de manera negativa en el uso de la -s genérica, que no se aplica a los compuestos del inglés. Esta influencia negativa del español tiende a decaer a medida que el nivel de los aprendientes aumenta.

Con la misma combinación de lenguas, González Alonso et al. (2016) investigaron el procesamiento de los compuestos deverbales en inglés por parte de diferentes grupos de bi- (L1 español-L2 inglés) o trilingües (L1 euskera-L2 español-L3 inglés y L1 español-L2 euskera-L3

inglés) y un grupo control de hablantes nativos de L1 inglés. Los autores incluyeron grupos de trilingües con una combinación espejo de las L1/L2 con el objetivo de explorar un posible efecto de la lengua dominante, algo que no se había tomado en cuenta en García Mayo (2006). En una tarea de decisión léxica cronometrada, que incluía compuestos deverbales atestiguados (p. ej., “*nutcracker*”, lit. “nuez-cascador”, “cascanueces”) o inventados (p. ej., “*nupscander*”) y sustantivos atestiguados (p. ej., “*people*”, “gente”) o inventados (p. ej., “*pundle*”), se observó una diferencia entre el grupo nativo y los grupos no nativos, tanto en los promedios de aciertos como en los tiempos medios de respuesta. En cuanto a las diferencias entre los grupos no nativos, los trilingües mostraron un tiempo de reacción significativamente mayor al de los bilingües. Los autores atribuyen este mayor tiempo de reacción a la conocida desventaja que tienen los bilingües en el tiempo de acceso al léxico en comparación con los monolingües (Lemhöfer et al., 2004; Gollan et al., 2008). Extendida a los trilingües, esta lógica explica la ligera desventaja de los trilingües frente a los bilingües del mismo modo que explica la diferencia entre los bi- y los monolingües. Con respecto al promedio de acierto, si bien el grupo de euskera L1 mostró un mayor acierto en la aceptación de los compuestos en inglés, esa diferencia entre los grupos no nativos no resultó significativa. Sin embargo, los autores observaron que los participantes dominantes de español (L1 español, bi- o trilingües) se destacaban más por sus aciertos en el rechazo de los no-compuestos que por la aceptación de los compuestos atestiguados. Los autores sugieren que este efecto podría estar relacionado con el hecho de que la forma de compuestos N+V+sufijo no existe en la lengua dominante de los participantes de español L1 y, por lo tanto, tienden a rechazarlos a primera vista. En cambio, los dominantes de euskera, que tienen una estructura similar en su lengua dominante la reconocen de manera más rápida. Además, dado que el euskera es una lengua aglutinante, estos participantes están más acostumbrados a procesar palabras complejas en su lengua dominante. Si bien este estudio no nos proporciona datos sobre los posibles errores de direccionalidad o de morfología de los compuestos, sí nos indica que, según la dominancia en las dos primeras lenguas, el procesamiento de los compuestos puede ser distinto en la L3.

3.5 Síntesis provisional

Como hemos visto hasta aquí, la composición nominal es un fenómeno complejo ampliamente estudiado tanto desde un punto de vista teórico (secciones 3.2 y 3.3) como desde el campo de la adquisición (sección 3.4).

De todos los estudios de adquisición comentados a lo largo de la sección 3.4, sobresalen algunos puntos. En L1, la edad de adquisición de los mecanismos de composición varía ligeramente en función de la lengua que se adquiere y, de manera indirecta, en función de la frecuencia de los compuestos en el *input* disponible (Clark & Hecht, 1982; Krott et al., 2009). A pesar de ello, a los 5 o 6 años de edad, todos los niños pueden producir y comprender los compuestos posibles en su lengua (Berman, 2009; Nicoladis, 2005).¹⁶¹ En adquisición bilingüe, si bien la coexistencia de las dos lenguas puede ser fuente de algunos errores, los niños perciben temprano la diferencia entre sus dos lenguas y son capaces de producir y comprender los compuestos de sus dos lenguas de manera independiente (Nicoladis, 1999, 2002b, 2002c, 2003, 2005). Los hablantes de herencia, que muestran resultados más variables en cuanto a sus aciertos con los compuestos, representan un caso particular. Es decir, en estos hablantes resulta difícil diferenciar si los mecanismos de composición han sido adquiridos de manera exitosa en la infancia, pero la posterior falta de *input* ha contribuido a debilitar los mecanismos de composición, o si simplemente no han tenido un *input* lo suficiente diversificado como para lograr establecer representaciones estables de los mecanismos de composición en su lengua de herencia (Garza-González, 2014). En el caso de la adquisición de L2, hemos observado una cierta tendencia a la influencia de la L1, sobre todo en cuanto a la **direccionalidad**, pero también en cuanto a la **morfología verbal y de número** en los compuestos deverbales (Desrochers et al., 2003; Lardiere & Schwartz, 1997; Liceras & Díaz, 2000b; Salomaa-Robertson, 2000; Slabakova, 2002). Es más, la morfología de número parece representar un desafío, aun cuando la L1 no debería ejercer influencia negativa (Murphy, 2000). También hemos notado que el **rasgo de animacidad** parecía tener cierto impacto con respecto a la productividad

¹⁶¹ Ninguno de los estudios sobre la adquisición de LN que hemos revisado versa exclusivamente sobre la comprensión de los compuestos, ya que se trata de un fenómeno poco explorado, particularmente en las lenguas de nuestro estudio. Nicoladis (2006) apunta que existen estudios a este respecto sobre el hebreo, una lengua en la que la producción de los compuestos es relativamente tardía. Los niños cuya L1 es el hebreo producen pocos compuestos antes de los cuatro años de edad, pero muestran una comprensión completa a los tres años (Berman & Clark, 1989; Clark & Berman, 1987). Nicoladis (2006) sugiere que la producción tardía podría relacionarse con la morfología compleja que presentan ciertas estructuras composicionales del hebreo. En cambio, en lenguas como el inglés, en la cual los compuestos se producen a una edad más temprana, hay menos evidencia de una brecha entre la comprensión y la producción de estas estructuras, a diferencia de lo que se suele documentar en otras áreas del lenguaje. Según Nicoladis, “these results suggest that comprehension and production of compound nouns emerge close together in development” (p. 115).

de los compuestos. Por una parte, los deverbales agentivos [+animado] parecen ser menos productivos que los instrumentales [-animado], tanto en francés como en español y, por otra, el rasgo [+animado] parece fomentar la percepción de unidad de los compuestos N+N (Liceras & Klassen, 2019). En L3, hemos visto que las dos lenguas previamente adquiridas podían influir en la adquisición de los mecanismos compositivos de la L3. una influencia posible entre lenguas tipológicamente más cercanas (francés-español) (Liceras et al., 2002; Pomerleau, 2000), pero también entre lenguas que se parecen solamente en cuanto a una estructura precisa (euskera-inglés) (García Mayo, 2006; González Alonso et al., 2016).

Como se indica en capítulos precedentes, este trabajo sigue la línea trazada por los estudios de adquisición de L2 y L3 sobre el tema y busca ampliarla añadiendo combinaciones de lenguas no estudiadas hasta la fecha, como es el caso del bilingüismo euskera-francés.

En lo referente a la composición, a causa de las posibilidades semejanzas (español-francés) y diferencias (español-euskera) que presenta el trío de lenguas de nuestro estudio –francés, euskera y español–, el análisis de este fenómeno resulta de gran utilidad para el estudio de la influencia interlingüística y para el campo de la adquisición de L3 de manera general. Como veremos más en detalle en la sección 3.6, el trío de lenguas de nuestro estudio nos permitirá explorar las posibilidades de transferencia del fenómeno en el español L3, la cual, en nuestro caso, puede darse o bien a partir del francés, que presenta la misma direccionalidad en los compuestos N+N y los compuestos deverbales, o bien a partir del euskera, que presenta la direccionalidad contraria al español.

3.6 Diseño experimental

El hilo conductor de nuestras preguntas de investigación es el análisis de la influencia de las dos primeras lenguas sobre el español L3 en nuestros aprendientes bilingües (francés-euskera). Como indicamos en el capítulo 1, en adquisición de una L3, si queremos hablar de influencia interlingüística, es particularmente importante explorar en profundidad las fuentes posibles de influencia. Con esto en mente, formulamos una primera pregunta preliminar (P1) que no concierne a la L3, sino las dos primeras lenguas. Contestar esta pregunta implica administrar a nuestros participantes las mismas pruebas en las tres lenguas del estudio, como veremos en la subsección 3.6.4. A partir de las respuestas obtenidas para esta pregunta preliminar, tendremos referencias

fiables para analizar las mismas estructuras en la L3 de nuestros participantes y plantearnos una serie de preguntas adicionales (P2-4):

Pregunta 1: ¿Qué características muestran los compuestos N+N y deverbales producidos por estos hablantes en sus dos primeras lenguas? ¿Son fieles a las descripciones que hemos podido observar en trabajos teóricos sobre los compuestos del francés y del euskera?

Pregunta 2: ¿Qué características muestran los compuestos N+N y deverbales producidos por estos hablantes de español L3? ¿Podemos observar alguna influencia de las lenguas previamente adquiridas en cuanto a la productividad y a la direccionalidad de estas construcciones?

Pregunta 3: ¿Tienen algún impacto los rasgos [+/-animado] y [+/-atestiguado] sobre la productividad de los compuestos N+N y deverbales en español L3? ¿Interactúan de algún modo estos rasgos con la influencia de las lenguas previamente adquiridas?

Pregunta 4: En cuanto a los compuestos deverbales, ¿presentan dificultades para estos aprendientes de L3 las características morfológicas de los compuestos deverbales del español, en particular la flexión verbal y la -s genérica?

3.6.1 Hipótesis

En cuanto a la pregunta 1, las hipótesis son bastante claras y corresponden esencialmente a lo que hemos descrito en las secciones 3.2.2 y 3.2.3. En francés, nos esperamos encontrar compuestos con núcleo inicial, tanto para los compuestos N+N como los deverbales y, en euskera, compuestos con el núcleo final para ambos tipos de compuestos. En cuanto a los N+N, dado que la tarea que administramos fomenta la producción de compuestos y no de otras estructuras (p. ej., N+de+N), los participantes deberían producir principalmente compuestos, tanto en francés como en euskera, a pesar de la menor productividad de esta estructura en francés. Si se observa una diferencia de productividad de los compuestos en cuanto al rasgo de animacidad para los N+N, el rasgo [+animado] debería favorecer la producción de compuestos, en ambas lenguas. Con respecto a los deverbales, si se observa una diferencia de productividad en cuanto a la animacidad, este efecto debería ocurrir al revés, es decir, el rasgo [+animado] debería inhibir la producción de deverbales, favoreciéndose otros mecanismos para aludir a agentes animados (p. ej. la derivación).

Finalmente, en cuanto a la forma de los compuestos deverbales en francés, deberíamos encontrar estructuras V+N, con un elemento verbal correspondiente a la conjugación de tercera persona del singular del presente de indicativo, y un complemento sin -s final. Con respecto al euskera, los compuestos deverbales deberían tener la estructura N+V+sufijo, con un sufijo agentivo

(-tzaile) para los compuestos [+animado], y un sufijo instrumental (-gailu principalmente, -tzeko en algunos casos) para los compuestos [-animado].

Al comprobar estas hipótesis sobre las dos primeras lenguas de nuestros participantes, podremos luego tener bases de comparación para explorar las hipótesis de influencia hacia la L3. En cambio, si algunas de las hipótesis planteadas sobre el euskera y el francés no se comprueban con nuestros datos, tendremos que ajustar las hipótesis sobre la influencia hacia la L3 que presentamos a continuación.

Según los modelos de transferencia de adquisición de L3 que hemos presentado en el capítulo 1, podemos postular una influencia o bien a partir del euskera, o bien a partir del francés. Según el *Typological Primacy Model* (TPM), el *Linguistic Proximity Model* (LPM) y el *Cumulative Enhancement Model* (CEM), la fuente de influencia debería ser el francés. En el trío de lenguas, el francés y el español son las dos lenguas tipológicamente más cercanas (TPM); los indicios (*micro-cues*) disponibles en las estructuras de los compuestos deberían llevar a los aprendientes a establecer una similitud estructural entre el francés y el español (LPM); la única influencia facilitadora o neutra puede proceder del francés (CEM). En cuanto al *L2 Status Factor*, la influencia debería proceder del francés solo si es la L2 de los participantes. En el caso contrario, se debería observar una influencia del euskera. Con respecto al llamado *L1 Factor*, la influencia debería proceder del francés solo si es la L1 de los participantes y del euskera en el caso contrario. Finalmente, el *Competition Model* (CM) predice una influencia principal de la lengua dominante, que es, en gran medida, una predicción similar a la del *L1 factor*. Sin embargo, en cuanto a la influencia potencial del euskera, es posible que la diferencia saliente entre la forma de los compuestos deverbales en euskera frente a los del español lleven a los aprendientes a un proceso de diferenciación que podría provocar una influencia del francés incluso si esta no es la lengua dominante.¹⁶²

Si aplicamos esta lógica a la influencia interlingüística en la formación de los compuestos, podemos establecer una serie de constataciones posibles según si la influencia procede del francés o del euskera. Agruparemos nuestras predicciones por tipo de compuesto, es decir, los N+N (N1-

¹⁶² Como vimos en la sección 1.4.6 del capítulo 1, cuando dos lenguas muestran divergencias importantes, el CM predice un proceso que se llama diferenciación. Volveremos sobre este aspecto en la discusión general del capítulo 4.

5) y los compuestos deverbales (D1-7) y las sufijaremos según si la fuente de influencia hipotética es el francés (FR) o el euskera (EU).

Como hemos visto a lo largo de este capítulo, la composición N+N es un proceso muy productivo en euskera (Azkarate, 1995; Coyos, 2004) y menos productivo en francés (Fradin, 2009) y en español (Kornfeld, 2009). Entonces, podríamos observar un efecto distinto según la fuente de influencia.

N1-FR: Todos los participantes que muestren una influencia del francés deberían producir más compuestos sintagmáticos preposicionales como “un barco de pirata”, descartando en muchos casos la opción del compuesto N+N, como “barco pirata”.

N1-EU: Todos los participantes que muestren una influencia del euskera deberían producir una mayoría de compuestos N+N y evitar los compuestos preposicionales u otras formas alternativas.

En cuanto a la direccionalidad de los compuestos N+N, considerando que el francés y el español comparten la misma direccionalidad –con el núcleo a la izquierda– y que el euskera presenta la direccionalidad inversa, deberíamos observar un efecto distinto según la fuente de influencia.

N2-FR: Los participantes que muestren una influencia del francés deberían producir compuestos N+N con la direccionalidad adecuada, es decir, con el núcleo inicial.

N2-EU: Los participantes que muestren una influencia del euskera deberían producir compuestos N+N con la direccionalidad inversa, es decir, con el núcleo final.

Con respecto a la animacidad, ya que no disponemos de datos específicos sobre la productividad de los N+N en función de ese rasgo en francés y en euskera, solo podemos suponer que lo comentado por Liceras y Klassen (2019) se aplica a todos los participantes de manera similar, es decir, si los compuestos inanimados muestran un mayor grado de separabilidad:

N3: Los participantes deberían producir una mayor tasa de compuestos N+N con los ítems cuyo referente es animado, ya que éstos muestran un menor grado de separabilidad.

El contraste entre los ítems cuya respuesta esperada está atestiguada en español y los ítems inventados podría también tener un impacto sobre la productividad y la direccionalidad de las respuestas proporcionadas, pero no se debería observar una influencia particular de las lenguas previamente adquiridas.

N4: Al tratarse de unidades léxicas presentes en la lengua y, por lo tanto, susceptibles de formar parte del *input* recibido por los aprendientes, los ítems atestiguados deberían elicitar una mayor tasa de respuestas esperadas (N+N con núcleo inicial) para todos los participantes.

En cuanto a la progresión a medida que los participantes avanzan en el proceso de adquisición, basándonos en estudios anteriores en AL2 (Liceras & Díaz, 2000b; Slabakova, 2002; entre otros), los aprendientes no deberían mostrar un efecto persistente de la influencia interlingüística sobre la direccionalidad, si es que la hay al inicio.

N5: Dado que todos los estímulos de nuestra tarea sirven para elicitar compuestos N+N bien formados en español (atestiguados o no), la tasa de producción de N+N no debería reducirse a pesar de la escasa productividad de esta estructura en español, pero la tasa de errores en cuanto a la direccionalidad sí debería disminuir.

En lo relativo a los compuestos deverbales, también establecemos una serie de hipótesis según si la influencia procede del francés o del euskera.

Considerando que los compuestos deverbales son productivos tanto en francés como en euskera, no podemos establecer una fuente de influencia clara para la hipótesis sobre la productividad de ese mecanismo en español.

D1: Todos los participantes deberían producir una mayoría de compuestos deverbales, ya que existe un mecanismo de composición equivalente en sus dos primeras lenguas. Las respuestas alternativas, como, por ejemplo, los compuestos preposicionales, deberían aparecer menos que en los datos sobre los N+N.

Considerando que el francés y el español comparten un proceso de formación de los compuestos deverbales similar (orden V+N) y que el euskera funciona de manera distinta (N-V-sufijo), deberíamos observar un efecto distinto según la fuente de influencia.

D2-FR: Los participantes que muestren una influencia del francés deberían producir compuestos deverbales con la direccionalidad adecuada, es decir, el elemento verbal primero, seguido por un sustantivo.

D2-EU: Los participantes que muestren una influencia del euskera deberían producir compuestos deverbales con la direccionalidad inversa, es decir, un sustantivo seguido por un verbal.

Sin embargo, el efecto de la animacidad podría afectar esta tasa de productividad, es decir, debería favorecer la producción de deverbales con los ítems inanimados en francés, favoreciéndose otras estructuras para los compuestos animados. En cuanto al euskera, no tenemos información

sobre la diferencia de productividad en función del rasgo, pero sí sabemos que no se usan los mismos procesos de derivación del verbo según se trate de un compuesto [+animado] o [-animado].

D3-FR: Los participantes que muestren cierta influencia del francés deberían producir menos compuestos deverbales con los ítems animados.

D3-EU: El rasgo de animación no debería influir sobre la productividad de los deverbales para los participantes que muestren cierta influencia del euskera. Si se observa un cambio en función de este rasgo, el cambio debería afectar la forma y no a la productividad como tal.

El contraste entre los ítems cuya respuesta esperada está atestiguada en español y los ítems inventados podría también tener un impacto sobre la productividad y la direccionalidad de las respuestas proporcionadas, pero no se debería observar una influencia concreta de las lenguas previamente adquiridas.

D4: Al tratarse de unidades léxicas presentes en la lengua y, por lo tanto, susceptibles de formar parte del *input* recibido por los aprendientes, los ítems atestiguados deberían elicitar una mayor tasa de respuestas esperadas (V+N) para todos los participantes.

En cuanto a la forma del elemento deverbal, ya que, en los compuestos deverbales del español, el primer elemento corresponde a la flexión verbal en tercera persona del singular del presente del indicativo y se usa la misma flexión en francés, el francés podría facilitar la producción de los deverbales en español. Sin embargo, dado que la flexión verbal no es un proceso bien asentado en los niveles iniciales de adquisición del español (Prévost & White, 2000), dicha influencia del francés podría ser difícil de percibir en los niveles iniciales. En cambio, considerando que el deverbal en euskera se forma a partir de un sufijo derivativo y no de una flexión verbal, como en las dos lenguas romances, las predicciones son distintas según la fuente de influencia.

D5-FR: Los participantes que muestren una influencia del francés deberían producir compuestos deverbales con un elemento verbal flexionado correctamente (en tercera persona del singular). Sin embargo, por lo menos en la primera recolección, es posible que aparezcan formas idiosincrásicas para el elemento deverbal; por ejemplo, puede aparecer el verbo en infinitivo, como “limpiarbotas”, o una flexión verbal alternativa, como “limpiobotas” (sobre todo la primera persona del singular siguiendo a Salomaa-Robertson, 2000).

D5-EU: Los participantes que muestren una influencia del euskera deberían derivar el deverbal a partir de un proceso similar a los deverbales del euskera, es decir, con un sufijo agentivo o instrumental, que podría corresponder a -dor(a) en español.

Finalmente, en cuanto a la -s que aparece al final del segundo elemento del deverbal, dado que ni el francés ni el euskera aplican esta -s genérica de manera sistemática, no podemos predecir una influencia de ninguna de las dos lenguas en particular, pero sí podemos predecir una dificultad para todos al inicio.

D6: Los participantes que muestren una influencia de cualquiera de sus dos primeras lenguas deberían, por lo menos al principio, producir los compuestos deverbales sin la -s final. Con una mayor experiencia con la lengua y un mayor *input*, la -s final debería aparecer en los datos de las segunda y tercera recolección.

En cuanto a la evolución longitudinal, sea cual sea la fuente de influencia inicial:

D7: Con una mayor experiencia y un mayor *input*, todos los participantes deberían converger hacia un mayor acierto en cuanto a la direccionalidad. El uso de formas verbales idiosincráticas debería decaer con el tiempo y una mayor exposición a la lengua, a medida que los aprendientes vayan dominando la flexión verbal en presente del indicativo. Sin embargo, en cuanto a la flexión verbal y el uso de la -s final, los estudios anteriores de L2 nos muestran que los errores persisten hasta niveles avanzados, por lo que la tasa de error debería decaer de manera más lenta y representar un mayor desafío que la direccionalidad.

3.6.2 Participantes

Los participantes de este estudio son los mismos que hemos presentado en el capítulo anterior, es decir, 52 alumnos de sexto grado de *collège* (10-11 años de edad) procedentes del País Vasco Norte, que asistían a un colegio cuya lengua de instrucción era el euskera, pero que contaba igualmente con 4-5 horas de cursos de francés a la semana.¹⁶³ Estos estudiantes eran también aprendientes de español como L3 y, al inicio de nuestro estudio, todos ellos se encontraban en las etapas iniciales de su proceso de adquisición (aproximadamente dos meses de enseñanza, 4 horas a la semana). A partir de un cuestionario de antecedentes lingüísticos completado por los participantes, pudimos dividir a nuestros participantes en los tres grupos siguientes, según la(s) lengua(s) hablada(s) en casa: francés y euskera (FR/EU), euskera (EU) o francés exclusivamente (FR). En los tres grupos experimentales, la edad de adquisición de la L2 o de la 2L1 variaba de un grupo a otro de manera significativa ($M=2.1$, $SD = 2.2$, $F(2,49) = 40.9$, $p < .001$), con un promedio particularmente alto en el grupo EU, lo cual indica que se trata de participantes que habían adquirido su L2, el francés, en un contexto más formal (en el jardín de infancia o al inicio de la

¹⁶³ Repetimos las características principales de nuestros participantes a modo de recordatorio, pero referimos al lector a la sección 2.6.2 del capítulo 2 para mayores detalles.

escuela primaria).¹⁶⁴ El grupo de control de nuestro estudio lo formaba un grupo de 21 niños de un colegio bilingüe inglés-español de la región de Valladolid (España) cuya lengua materna y dominante era el español. La distribución e información general sobre nuestros participantes aparece en la Tabla 3.6, la división por grupos es la misma que la que hemos presentado en el capítulo anterior.

Tabla 3.6. Resumen del perfil de los participantes según la(s) lengua(s) hablada(s) en casa.

	ES (n=21)	EU (n=14)	FR (n=15)	FR/EU (n=23)
Edad	11.4 (0.2)	11.3 (0.3)	11.4 (0.3)	11.3 (0.3)
Edad de adquisición L2 o 2L1	-	4.1 (1.8)	2.5 (1.8)	0.6 (1.3)
Horas de enseñanza en español	-	28 (0)	28 (0)	28 (0)
Prueba de nivel en español	117.6 (2.6)	12.8 (10.8)	12.5 (7.9)	9.8 (5.3)
Prueba de nivel en euskera	-	109.9 (8)	104.9 (6.8)	100.8 (11.1)
Prueba de nivel en francés	-	109.6 (10.5)	112.3 (4)	108 (7.9)

Nota. Los datos representan el valor promedio acompañado de la desviación estándar entre paréntesis. Los resultados de las pruebas de nivel son sobre 120.

Recordamos al lector que los resultados bajos a la prueba de nivel en español son coherentes con la cantidad reducida de horas de enseñanza que habían recibido los participantes (28 horas) al empezar el estudio longitudinal.

3.6.3 Instrumentos

Hemos utilizado los mismos instrumentos que los presentados en el capítulo 2. El primer instrumento que utilizamos para obtener información sobre nuestros participantes es un cuestionario sobre su perfil sociolingüístico (completado por uno de los padres de los participantes). A partir de esta información pudimos establecer cuáles eran los participantes correspondientes al perfil que buscábamos.

Para completar el perfil de los participantes, utilizamos una prueba de nivel de repetición elicitada en las tres lenguas de los participantes y una prueba de memoria a corto plazo (de

¹⁶⁴ Para referirnos a las lenguas previas de nuestros participantes, en los casos de bilingüismo secuencial (grupos EU y FR), distinguimos entre L1 y L2 y, en el caso de bilingüismo simultáneo (grupo FR/EU), hablamos de 2L1.

repetición de dígitos) en la lengua dominante de cada participante. Además, una tarea escrita de denominación de imágenes nos permitió evaluar de manera experimental el desarrollo de la formación de los compuestos en las tres lenguas del estudio. A continuación, presentamos la tarea de denominación de imágenes, y referimos el lector a la sección 2.6.3 del capítulo 2 para los detalles sobre los demás instrumentos.

3.6.3.1 Tarea de denominación de imágenes

Para estudiar la formación de los compuestos, utilizamos el instrumento de denominación de imágenes descrito en la sección 2.6.3.3 del capítulo 2, que incluía, por un lado, ítems para elicitación de la concordancia de género y, por otro, ítems para elicitación de compuestos. Los primeros fueron analizados en el capítulo 2 y, a continuación, nos concentramos en los segundos. De los 68 ítems de la prueba, 32 están relacionados con la formación de compuestos (16 para los N+N y 16 para los deverbales). La lista completa de ítems aparece en los Anexos 4 y 5. A los 68 ítems se suman 6 ítems de práctica (2 para cada tipo de compuesto y 2 para el género) para asegurarnos de que los participantes entendían bien la prueba y el tipo de respuesta esperada. Estos ítems se inspiran en dos fuentes utilizadas en varios estudios anteriores para elicitación de compuestos: la imagen (Clark et al., 1986; Clark & Hecht, 1982; Lardiere & Schwartz, 1997) y la paráfrasis (Gordon, 1985; Murphy, 2000; Salomaa-Robertson, 2000; entre otros). Si bien esas fuentes se pueden usar en tareas distintas, para elicitación de los deverbales hemos considerado necesario el uso de ambas para evitar respuestas en blanco, dado el bajo nivel de nuestros participantes, especialmente durante la primera recogida de datos.

La Figura 3.1 ilustra un ítem de cada tipo, el primero relacionado con la formación de un compuesto N+N y el segundo con la formación de un compuesto verbal. Así pues, para asegurar la mayor semejanza posible entre los dos tipos de ítems, todos los estímulos incluían una imagen y un breve texto de acompañamiento.



Figura 3.1. Ítems 59 y 52 de la prueba de denominación de imágenes en español.

En cada ítem, el participante debía observar la imagen, leer una oración del tipo “Identifica esto en pocas palabras” (para los NN) o “Quiero romper 100 botellas, necesito un ...” (para los compuestos deverbales) y proporcionar una respuesta adecuada. La respuesta esperada para el ítem 59 que aparece en la Figura 3.1 sería “un gato policía” y para el 52 “un rompebotellas”. Para evitar respuestas en blanco, los participantes tenían acceso a un mini-diccionario visual de dos páginas (ver Anexo 2) que contenía todos los sustantivos necesarios para formar los compuestos (lo único que no tenían eran ilustraciones de los verbos, pero las propias imágenes de los estímulos eran lo suficientemente claras como para que los verbos se pudieran entender únicamente con la ayuda del texto y de la imagen). Para evitar recurrir a las otras lenguas conocidas por el participante, optamos por un mini-diccionario visual y no por uno multilingüe como en Licerias et al. (2002).

Las variables contempladas en la elaboración de los estímulos se inspiraban en estudios anteriores sobre el fenómeno. Para los compuestos N+N, consideramos el género (masculino o femenino), si están atestiguados o no en español y si hacían referencia a entidades animadas o inanimadas. Para los deverbales, las variables eran las mismas, salvo la de género, dado que, de manera general, los compuestos deverbales son masculinos en español.

La variable género ha sido elegida principalmente para poder contrastar la concordancia de género en esos ítems frente a los ítems analizados en el capítulo 2, por lo que no nos detendremos en este criterio, ya que no hay nada específico en la literatura que nos lleve a pensar que el género pueda tener un efecto sobre la direccionalidad o la productividad de los compuestos N+N.

En cuanto al hecho de si los compuestos elegidos están atestiguados o no, hemos decidido incluir tanto ítems atestiguados como inventados, ya que la frecuencia y la productividad de una forma dada suele jugar un papel en la adquisición en la L1 (Clark, 2009). Sin embargo, dado el bajo nivel de español de nuestros grupos experimentales y el *input* reducido que han recibido, es probable que ese factor no sea relevante, por lo menos en las etapas iniciales de adquisición. Tanto en el caso de los N+N como en el de los deverbales, hemos comprobado si los compuestos estaban atestiguados o no mediante la consulta del *Diccionario de la Lengua Española 23^a ed.* (DLE; Real Academia Española [RAE], 2014) y del corpus “*Spanish 2009*” de *Google NGram* (Lin et al.,

2012).¹⁶⁵ Varios de los ítems han sido inspirados o tomados de estudios anteriores sobre los compuestos (Liceras et al., 2002, 2004; Liceras & Díaz, 2000b). Todos los compuestos inventados se asemejan a compuestos existentes y respetan los patrones de formación de la lengua (por ejemplo, “gato policía” ~ “perro policía”, “lanzapasteles” ~ “lanzallamas”, etc.). Si se observa un efecto con los compuestos atestiguados frente a los noveles, es probable que, dentro de los compuestos atestiguados, ese efecto varíe según la frecuencia de los compuestos en cuestión. Por ejemplo, es probable que algunos participantes hayan oído alguna vez “hombre araña”, pero nunca hayan oído “abeja reina”, dos compuestos atestiguados según el *DLE* (RAE, 2014). Por lo tanto, a la hora de analizar nuestros datos, será necesario tomar en cuenta la frecuencia de esos compuestos y no solo el valor binario atestiguado/inventado.

En cuanto al rasgo de animacidad, si bien no ha sido incluido a menudo en los estudios anteriores sobre los compuestos en AL2, como hemos visto en la sección 3.4.3, Liceras y Klassen (2019, p. 8) afirman que podría tener un impacto sobre la percepción de separabilidad de los dos elementos del compuesto, por lo que lo hemos incluido en nuestro diseño. Además, como hemos visto en la descripción de las tres lenguas de nuestro estudio –sobre todo en francés y en español– se trata de un rasgo que tiene un impacto sobre la productividad de ciertos tipos de compuesto, como los deverbales. La Tabla 3.7 muestra la repartición de los ítems N+N según estos tres criterios (número de ítems entre paréntesis), así como un ejemplo para cada combinación.

¹⁶⁵ Somos conscientes de la arbitrariedad relativa del criterio atestiguado/inventado. Dado que el proceso de composición sigue siendo productivo, se puede buscar en Google cualquiera de los compuestos que hemos elegido y encontrar resultados. Por esa razón hemos preferido consultar fuentes editadas (el *Google NGram* contiene el texto de casi un millón de libros en español) para atestiguar del uso de los compuestos elegidos. Tres de los compuestos N+N que consideramos como atestiguados (“barco pirata”, “mujer araña” y “carta bomba”) no aparecían en el *DLE* (RAE, 2014), pero sí aparecían en el corpus *Google NGram* (Lin et al., 2012). Todos los N+N no atestiguados (salvo “hombre hormiga”) no aparecían ni en el *Google NGram*, ni en el *DLE*. En cuanto a los deverbales, todos los compuestos que consideramos como atestiguados aparecen en el *DLE* y en el *NGram* salvo “pelapatatas” que solo aparece bajo la forma “pelapapas” en el *Google NGram*. Los deverbales no atestiguados no aparecen ni en el *DLE*, ni en el *Google NGram*, salvo “abrepuestas” que sí aparece en el *Google NGram*. En resumen, somos conscientes de dos discrepancias en nuestra clasificación: “hombre hormiga” y “abrepuestas” y tomamos nota de ello a la hora de analizar los datos.

Tabla 3.7. Repartición de los ítems N+N según los criterios de género, animacidad y si están atestiguados o no.

Género	Atestiguado/Inventado	Animacidad	Ejemplo
M (8)	A (4)	A (2)	Un perro policía
		I (2)	Un barco pirata
	N (4)	A (2)	Un hombre hormiga
		I (2)	Un bolígrafo pirata
F (8)	A (4)	A (2)	Una mujer araña
		I (2)	Una carta bomba
	N (4)	A (2)	Una mujer hormiga
		I (2)	Una silla pirata

En el caso de los deverbales, la animacidad se representa en el contraste agentivo/instrumental y cobra aquí más importancia, ya que, en una de las lenguas de nuestros participantes (el euskera), el proceso formación de los compuestos deverbales difiere según si estos son agentivos o instrumentales (Azkarate, 1995). Por otro lado, en francés y en español, los compuestos deverbales suelen hacer referencia más a menudo a entidades inanimadas (instrumentos). La Tabla 3.8 muestra la repartición de los ítems para elicitación de compuestos deverbales según los criterios mencionados (número de ítem entre paréntesis) y un ejemplo para cada combinación.

Tabla 3.8. Repartición de compuestos deverbales según su animacidad y si están atestiguados o no.

Atestiguado/Inventado	Animacidad	Ejemplo
Atestiguado (8)	Agentivo (4)	Un limpiabotas
	Instrumental (4)	Un abrebotellas
Inventado (8)	Agentivo (4)	Un tragateléfonos
	Instrumental (4)	Un lanzapasteles

Dado que la tarea tiene también versiones en euskera, en francés hemos tenido que validar que cada ítem seguía los mismos patrones de formación en cada idioma, es decir, si tenemos el hipotético “lanzapasteles” en español, también podríamos tener los equivalentes “*bixkotx jaurtigailu*” (euskera) y “*lance-gâteau*” (francés).¹⁶⁶ Para estas versiones en otros idiomas, hemos

¹⁶⁶ Sin embargo, solo hemos comprobado si un compuesto está atestiguado o no en español, no en euskera o en francés. Aplicar el criterio a las tres lenguas hubiera añadido demasiada complejidad al estudio, tanto en la preparación de los

traducido las oraciones que aparecían en las diapositivas de los compuestos deverbales para adaptarlas al compuesto esperado. Por ejemplo, la oración que aparece con el ítem 52 en euskera incluye el verbo *hautsi* ‘romper’ y el sustantivo *botila* ‘botella’ para obtener *botila hauskailu* ‘rompebotellas’, mientras que en francés incluye el verbo *briser* ‘romper’ y el sustantivo *bouteille* ‘botella’. Todos los ítems fueron cuidadosamente seleccionados de manera que pudieran elicitarse un compuesto deverbal o un N+N en los tres idiomas del estudio.

3.6.4 Protocolo de administración de las tareas

Referimos el lector a la sección 2.6.4 del capítulo 2 para más información sobre el protocolo de administración de las tareas que sirvieron para clasificar a los participantes (el cuestionario de perfil lingüístico, la prueba de nivel y la narración oral).

En cuanto a la tarea de descripciones de imágenes, esta se administró en los tres idiomas de los participantes (empezando con el español, la L3), con un intervalo mínimo de una semana entre cada idioma. Las pruebas en las tres lenguas se administraron una sola vez, después de 28 horas de enseñanza en español, y la versión en español se volvió a administrar después de 140 horas de enseñanza en español y una vez más después de alrededor de 200 horas, para ver cómo evolucionaba el dominio del fenómeno.

En una clase equipada con un proyector, todos los participantes realizaron la tarea simultáneamente y de manera individual. Antes de empezar, el investigador explicó la tarea a los participantes y les entregó una hoja de respuesta individual y el mini-diccionario visual (ver Anexo 2). Las explicaciones se dieron en el idioma de la tarea, pero dado que los participantes tenían un nivel bajo de dominio del español, el investigador y el docente contestaron a las preguntas en francés o en euskera cuando resultó necesario para aclarar las instrucciones de la tarea. Después, el investigador prosiguió con los seis ítems de práctica y escribió las respuestas en la pizarra. Antes de empezar, el investigador explicó que cada ítem aparecería durante solo 25 segundos en la pantalla y que, por lo tanto, tenían que escribir la primera respuesta que les viniera a la mente. Se insistió también en que las respuestas eran individuales y en que era importante escribir una respuesta para cada diapositiva. A la mitad de la tarea, se realizó una pausa de 2 minutos. La tarea

estímulos como en la recogida de datos. Se trata, entonces, de una tarea adaptada a las dos primeras lenguas de los participantes con modificaciones mínimas.

desde el inicio de las explicaciones hasta el final duró entre 35 y 45 minutos, según la duración de las explicaciones.

3.7 Resultados

Antes de comprobar nuestras hipótesis con los datos de español como L3, empezaremos con la presentación de nuestros datos de control sobre el euskera y el francés para contestar a nuestra primera pregunta de investigación, es decir, para asegurarnos de que las respuestas de nuestros participantes en sus dos primeras lenguas se corresponden con la descripción del fenómeno que hemos presentado en la sección 3.2. Ajustaremos nuestras hipótesis conforme a lo encontrado en nuestros datos de euskera y de francés. Además, para comprobar la adecuación de la tarea que nos sirvió para elicitación los diferentes tipos de compuestos, presentaremos los resultados de nuestro grupo de control de español antes de presentar los resultados de español L3. Si bien hay muchos paralelos entre las estructuras de los compuestos N+N y los deverbales, podría resultar abrumador presentar los datos de estas dos formas de manera conjunta. Por lo tanto, tras una presentación de nuestra metodología general de análisis, empezaremos con los resultados sobre los compuestos N+N, continuaremos con los compuestos deverbales y terminaremos con una reflexión global sobre los dos tipos de compuestos.

Las respuestas que presentamos a continuación proceden de la tarea escrita de denominación de imágenes y fueron codificadas de la misma manera para las tres lenguas del estudio. Para los NN, consideramos 1) el tipo de respuesta producida (esperada u otra) y 2) la direccionalidad (núcleo inicial o núcleo final). En el caso de los deverbales, consideramos 3) el tipo de respuesta producida (esperada u otra), 4) la direccionalidad (V+N o N+V), 5) la flexión verbal (acierto o error) y 6) la -s final (acierto o error).¹⁶⁷ Estas seis respuestas binarias fueron analizadas, con el programa *R* (R Core Team, 2018), que utilizamos para construir varios modelos de regresión logística de efectos mixtos mediante el paquete *lme4* (D. Bates et al., 2015), como hicimos en el análisis de los datos sobre el género presentados en el capítulo 2. Para comodidad del lector, repetimos a continuación los pasos que seguimos para construir nuestros modelos. En todos los casos, incluimos las estructuras aleatorias máximas, salvo cuando el nivel de complejidad del modelo impedía su convergencia. Cuando los modelos no convergían, eliminamos, en primer

¹⁶⁷ En el caso del euskera, los dos últimos criterios no se aplican.

lugar, la correlación entre las pendientes (*slopes*) y los interceptos (*intercept*) aleatorios, las interacciones en las pendientes aleatorias, luego las pendientes aleatorias y, por último, las interacciones de mayor nivel entre los efectos fijos, conservando, como mínimo los interceptos de sujeto e ítem. La inclusión de estos efectos aleatorios, como recomiendan Baayen et al. (2008), permite tomar en cuenta la variabilidad aleatoria en nuestros datos experimentales. Para cada respuesta, tras el ajuste de un primer modelo que convergía, simplificamos el modelo eliminando los efectos uno por uno, guiándonos por el resultado de la función *dredge* de *MuMIn* (Bartón, 2019), para seleccionar el modelo que tenía la mejor bondad de ajuste, según su AIC (*Akaike Information Criterion*). Por lo tanto, para cada respuesta, el modelo resultante solo incluía los efectos que mejoraban las predicciones globales del modelo, sin aumentar innecesariamente su complejidad.

3.7.1 Los N+N

3.7.1.1 Euskera L1/2L1/L2

Recordamos al lector que, según la literatura revisada sobre los compuestos en euskera, la gran mayoría de los N+N deberían tener un núcleo final, como en inglés. Además, la composición de tipo N+N sigue siendo productiva en euskera y debería ser la opción privilegiada por nuestros participantes. Sin embargo, en nuestros datos, de todas las respuestas ($n=832$), un 77.64 % son de tipo N+N, mientras que apenas un 0.36 % son respuestas en blanco y las demás corresponden a diferentes estructuras inesperadas. De estas últimas, encontramos un 10.46 % de respuestas que corresponden a la estructura N + posposición + N, como en (3.124), un 1.80 % que corresponde a N + N + posposición, como en (3.125), un 3,24 % de sustantivos solos, un 3,72 % de sustantivos modificados por un adjetivo y, finalmente, un 2,76 % de respuestas de otras estructuras, como en (3.126). En la categoría “otras estructuras” hemos agrupados estructuras coordinadas, como (3.126), pero también estructuras que contenían un sustantivo con un sufijo derivativo que cambiaba la clase de palabra a un participio, como en (3.126).¹⁶⁸

¹⁶⁸ La mayoría de las abreviaturas que aparecen en las glosas también aparecen en la sección 3.2.3. A modo de recordatorio, la abreviatura ADN hace referencia a “adnominal”, un tipo de posposición documentado en las gramáticas y diccionarios del euskera (de Rijk, 2008; Michelena, 1987). Esta posposición se usa de manera similar a la preposición

- (3.124) a. *itsaso.ko izarra* / *itsas izarra*
mar.ADN estrella / mar estrella
“estrella de mar” / “estrellamar”
- b. *pirat.en kadera* / *pirata kadera*
pirata.GEN-PL silla / pirata silla
“silla de los piratas” / “silla pirata”
- c. *pirata.ko kadera* / *pirata kadera*
pirata.ADN silla / pirata silla
“silla del pirata” / “silla pirata”
- (3.125) a. *petarr.ak gutun.ean* / *bonba gutuna*
bomba.DET-PL carta.LOC-SG / bomba carta
“bombas en la carta” / “carta bomba”
- b. **neska ur.eko.a* / *olagarro emaztea*
chica mar.ADN.DET-SG / pulpo mujer
“chica del mar” / “mujer pulpo”
- c. *#auto.a petarr.ekin* / *bonba auto*
auto.DET-SG bombas.SOC-PL / bomba auto
“el coche con bombas” / “coche bomba”
- (3.126) a. *auto urdin.a eta bomba* / *bonba auto*
auto azul.DET-SG y bomba.DET-SG / bomba auto
“coche azul y bomba” / “coche bomba”
- b. *#auto bonbatua* / *bonba auto*
auto bombear.PTCP / bomba auto
“coche bombardeado” / coche bomba

Este primer acercamiento a los datos nos revela que los participantes utilizan una variedad de mecanismos alternativos, si bien usan mayoritariamente la composición de tipo N+N. Entonces, a pesar de ser productiva, la composición N+N no representa la única opción para modificar a un sustantivo. A este respecto, la Figura 3.2 muestra la proporción de respuestas de cada tipo según el ítem. Las etiquetas corresponden a la respuesta esperada para cada ítem, con su equivalente en español entre paréntesis. Las etiquetas en mayúsculas corresponden al rasgo [-animado] y las minúsculas al rasgo [+animado].

“de” en español, es decir, cumple una función de nexos entre dos palabras, sin necesariamente conllevar un significado propio.

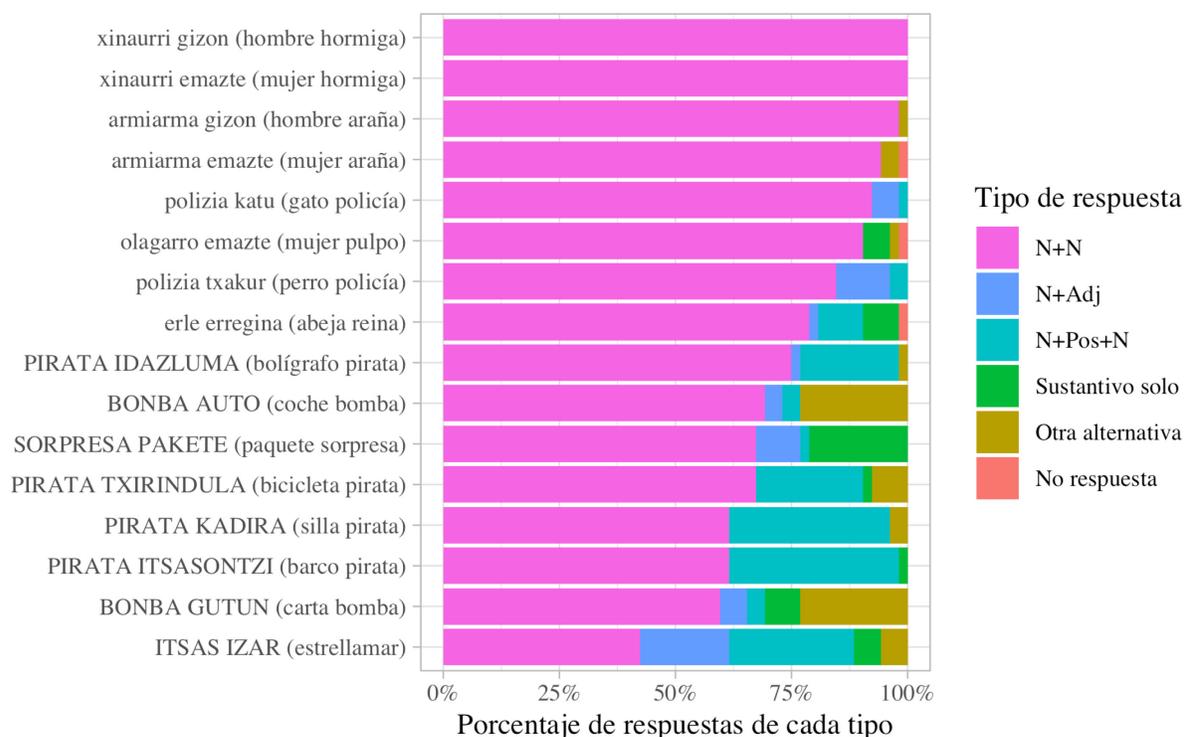


Figura 3.2. Repartición de los tipos de respuestas según los ítems N+N en euskera.

A primera vista, constatamos que el rasgo de animacidad parece estar vinculado con el tipo de respuesta elicitada. Todos los compuestos animados aparecen en la parte superior del gráfico, es decir, con una mayor proporción de compuestos N+N, mientras que los inanimados aparecen en la parte inferior. Para evaluar los efectos de animacidad y de grupos sobre el tipo de respuesta (1=N+N; 0=Otros), construimos un primer modelo de regresión logística de efectos mixtos, en el que incluimos estas variables como efectos fijos y también la posible interacción animacidad \times grupo. En cuanto a la estructura de los efectos aleatorios, incluimos un intercepto aleatorio por participante e ítem y una pendiente aleatoria por participante para la animacidad. Los factores fueron codificados de la siguiente manera (el nivel de referencia corresponde al primer nivel de codificación): animacidad (-.5=inanimado, .5=animado), grupo (1=EU, 2=FR, 3=FR/EU). Dado que ni el factor grupo ni su interacción con el factor animacidad mejoraban la bondad de ajuste del modelo (según el AIC), nuestro modelo final solo incluye el factor fijo de animacidad que resulta muy significativo ($\beta = 2.2022$, $SE = 0.4959$, $z = 4.441$, $p < .001$). El valor positivo de este coeficiente nos indica que el rasgo [+animado] favorece la producción de compuestos N+N, lo que

apunta, como lo habíamos planteado, a un mayor grado de inseparabilidad de los compuestos animados.

Para continuar el análisis, nos enfocamos en las respuestas que se acercan a lo que podríamos considerar como un compuesto N+N o un compuesto posposicional (que sería el equivalente de la estructura N+prep+N en francés y en español) para evaluar la direccionalidad. El análisis de esos datos resulta sorprendente, ya que encontramos una mayoría de respuestas (57.92 %) con el núcleo inicial. Sin embargo, dentro de esos diferentes tipos de respuestas, se encuentra mucha variación. Por un lado, tenemos los sintagmas N+Pos+N que respetan el orden del núcleo final característico del euskera en un 90.67 % y, al otro extremo, los N+N+Pos que tienen un núcleo inicial en un 76.92 % de los casos. Los compuestos N+N se encuentran cerca del promedio global con un 64.62 % de núcleo inicial. Los promedios globales no nos proporcionan detalles sobre la variación de un ítem a otro en cuanto a la direccionalidad; la Figura 3.3 ilustra el efecto de los ítems de manera más clara.

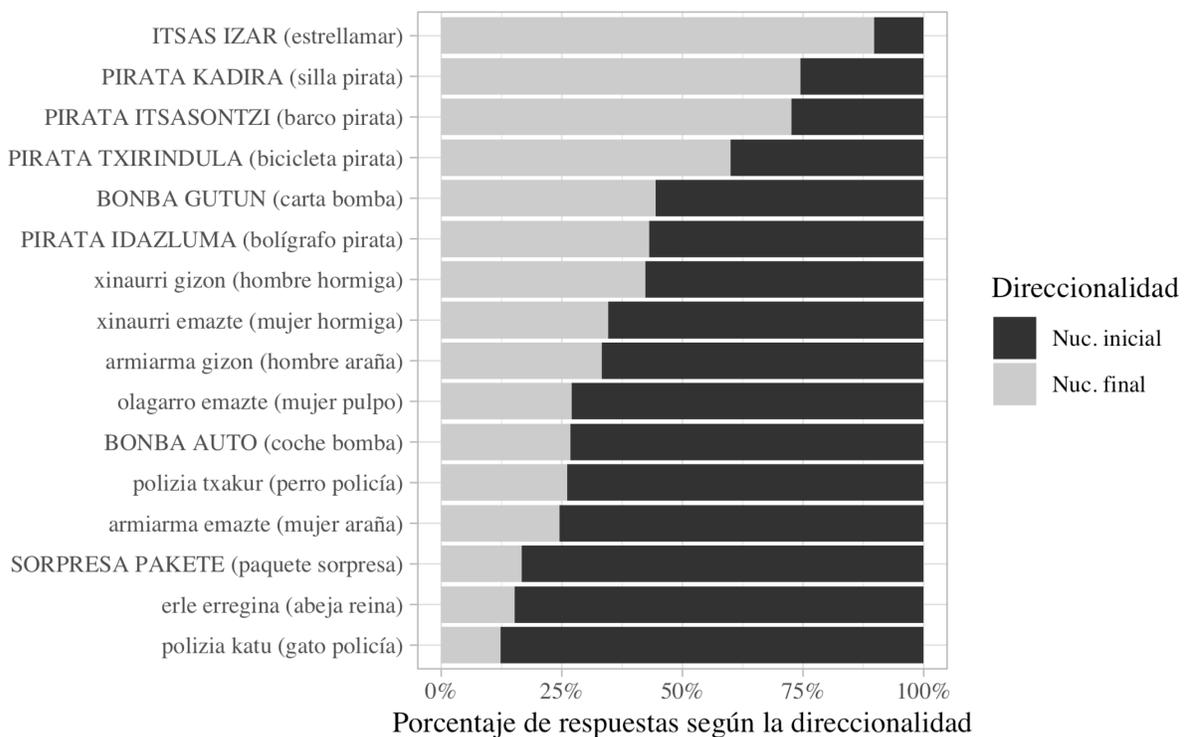


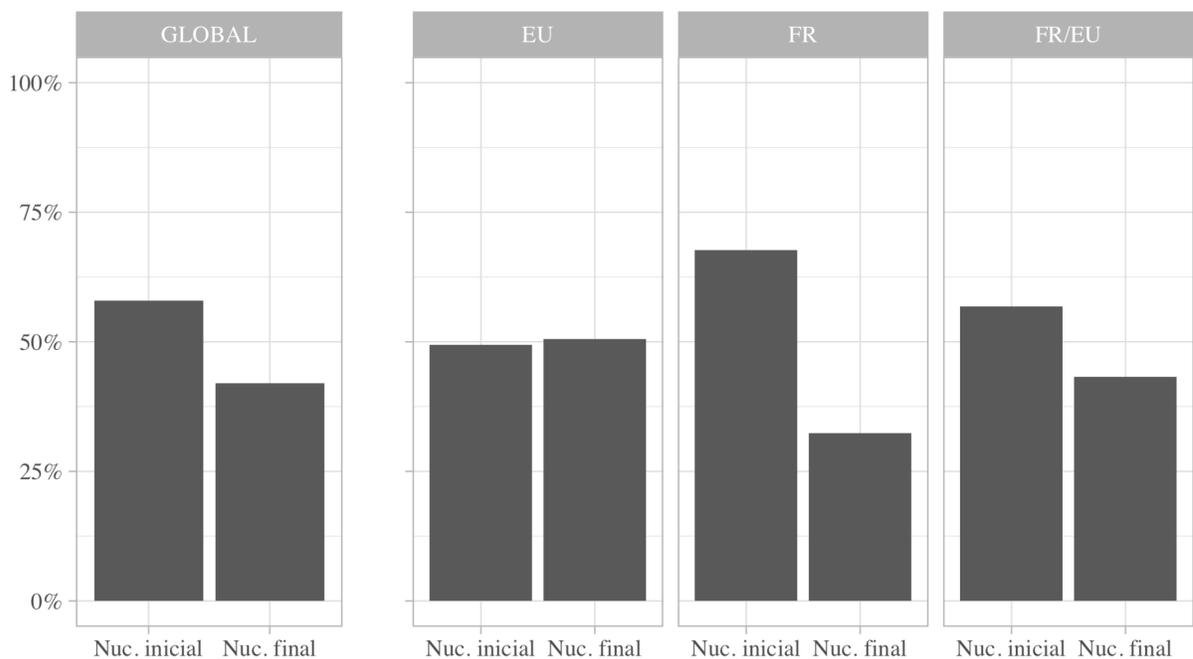
Figura 3.3. Repartición de la direccionalidad de las respuestas N+N en euskera según los ítems.

Antes de continuar, cabe mencionar que el compuesto “*erle erregina*” (“abeja reina”) representa una excepción a la regla general del euskera en cuanto a la posición del núcleo, si lo

interpretamos como su equivalente en español. Así, si consideramos que el núcleo es “abeja” ya que “una abeja reina es un tipo de abeja” (prueba 4 de Arnaud 2003, pp. 8-9), el núcleo “*erle*” (“abeja”) se ubica a la izquierda del compuesto “*erle erregina*” en euskera. Por otro lado, también se podría considerar que el núcleo es “reina” si consideramos que “una abeja reina es un tipo de reina”. Ya que se trata de un compuesto con un núcleo ambiguo y que podría distorsionar la interpretación de nuestros datos, lo quitamos de nuestro análisis del euskera.¹⁶⁹ Esta nueva consideración permite disminuir la tasa de compuestos con el núcleo inicial a un 56.61 %.

Con esa tasa, todavía estamos lejos de lo que deberíamos esperarnos de una lengua que, en teoría, favorece el núcleo final en los compuestos N+N. Esta desviación de la norma parece explicarse por un factor principal: la lengua hablada en casa, como aparece en la Figura 3.4. El grupo FR muestra una tasa más elevada de compuestos con el núcleo inicial.

Direccionalidad de los compuestos, de manera global y por grupo



¹⁶⁹ Cabe mencionar que, en euskera, “*erle erregina*” sigue el mismo patrón que “*erle langile*” (lit. “abeja obradora”, es decir, “abeja obrera”). En este caso, se podría considerar que “*langile*” cumple una función adjetival y estaría en la posición adecuada. También podríamos considerar que el núcleo es “*langile*” si consideramos que, semánticamente, la función de obrera es más importante que el aspecto físico de abeja. En fin, “*erle erregina*” y “*erle langile*” forman parte de los compuestos difíciles de analizar en euskera, y a veces la frontera entre núcleo y modificador es borrosa. Veremos más adelante que este difícil análisis se aplica también a este ítem en francés.

Figura 3.4. Direccionalidad de los compuestos N+N en euskera, según los grupos lingüísticos.

Para comprobarlo, realizamos una regresión logística de efectos mixtos para la direccionalidad (0=Núcleo final, 1=Núcleo inicial), partiendo de la misma estructura de factores (fijos y aleatorios) que en la regresión anterior sobre la productividad. La interacción grupo \times animacidad no mejoraba la bondad del ajuste del modelo, por lo que nuestro modelo final solo incluye los efectos fijos grupo y animacidad, por separado. El efecto de animacidad es significativo ($\beta = 1.6351$, $SE = 0.5704$, $z = 2.867$, $p = .004$), es decir que el rasgo [+animado] contribuye a la producción de compuestos con el núcleo inicial, la direccionalidad inversa para el euskera. En cuanto a la diferencia entre los grupos, una prueba *posthoc* de comparaciones múltiples de medias (Contraste Tukey) realizada con el paquete *multcomp* (Hothorn et al., 2008) nos revela que son significativos los contrastes FR - EU ($\beta = 1.0806$, $SE = 0.4276$, $z = 2.527$, $p = 0.012$) y FR/EU - FR ($\beta = -0.8741$, $SE = 0.3709$, $z = -2.357$, $p = .018$), pero que el contraste FR/EU - EU no lo es ($\beta = 0.2065$, $SE = 0.3923$, $z = 0.526$, $p = .599$).¹⁷⁰

Si bien el factor grupo explica una parte de nuestros datos, aun considerando solo los datos del grupo EU, los resultados no se parecen a lo que se esperaba según la descripción del fenómeno en la literatura. En cuanto al factor de animacidad, parece coincidir en gran medida, con una excepción que no fue tomada en cuenta en nuestro diseño experimental para el euskera. Efectivamente, Coyos (2004, p. 58) afirma que ciertos compuestos poco frecuentes se forman con el núcleo a la izquierda. En estos compuestos, se considera que el no-núcleo actúa como un atributo del núcleo, una función que Coyos relaciona con la del adjetivo calificativo. Si consideramos el no-núcleo como un adjetivo, entonces, estos compuestos respetan el orden canónico del euskera. Reproducimos en (3.127) un ejemplo de Coyos (2004, p. 58) para ilustrar el fenómeno.

(3.127) *etxe- orratz*
 casa aguja
 “rascacielos”

¹⁷⁰ Para comparar los efectos de nuestras variables categoriales de más de dos niveles (como los grupos), usaremos siempre el test *posthoc* de *Multiple Comparisons of Means: Tukey Contrasts* del paquete *multcomp* (Hothorn et al., 2008).

En (3.127), el no-núcleo caracteriza la forma del edificio y, por lo tanto, aparece en la posición del adjetivo. Sin embargo, en el diccionario, también encontramos contraejemplos en los cuales la relación parece ser atributiva, pero que presentan un núcleo final, como en (3.128). En este ejemplo, la “aguja” caracteriza la forma del tacón, de manera similar a lo que ocurre en el ejemplo (3.127). Entonces, parece existir cierta variabilidad en cuanto al orden de los componentes en estos compuestos atributivos en euskera. En cambio, en (3.129), la relación entre “*orratz*” y el otro elemento correspondería a lo que Arnaud (2016) llama un compuesto subordinativo relacional, es decir que, existe una relación entre “*buru*” y “*orratz*” que se puede parafrasear como “un N1-N2 es un N2 que tiene un N1” (“un alfiler, es una aguja que tiene una cabeza”).

- (3.128) *orratz- orpo*
 aguja talón
 “tacón de aguja”
- (3.129) *buru- orratz*
 cabeza aguja
 “alfiler”

Los compuestos relacionales como el ejemplo (3.129) no parecen presentar ambigüedades en cuanto a la direccionalidad. Desafortunadamente, la mayoría de los ítems de nuestra prueba pueden ser interpretados más bien como compuestos atributivos. A este respecto, reproducimos en (3.130) un compuesto aún más cercano a los ítems [+animado] de nuestra tarea.

- (3.130) a. *giz. otso*
 hombre lobo
 “hombre lobo”
- b. *otso- gizon*
lobo- hombre
 “hombre lobo”

Estamos ante un caso en el que las dos formas son posibles, con la direccionalidad opuesta, pero con un mismo significado. Se trata de un ejemplo similar a (3.127) y (3.128), pero esta vez con los dos mismos elementos invertidos. El compuesto en (3.130) se parece a siete de los ocho ítems animados de nuestra tarea, es decir, se puede interpretar como “un N1-N2 es un N1 que se parece a un N2”.¹⁷¹ Con estos siete ítems, la tasa de núcleo a la izquierda alcanza los 71.18 %. En

¹⁷¹ Los ítems en cuestión son: hombre hormiga, mujer hormiga, hombre araña, mujer araña, mujer pulpo, gato policía y perro policía. Si bien “perro policía” se puede interpretar como “un perro adiestrado que ayuda a la policía” y no

cambio, en los ítems inanimados de nuestra tarea, las relaciones eran más diversas, pero cuatro ítems destacan con una mayoría de respuestas con el núcleo a la derecha, la direccionalidad esperada: “itsas izar” (“estrellamar”, 89,74 %), “pirata kadirra” (“silla pirata”, 74,50 %), “pirata itsasontzia” (“barco pirata”, 72,54 %) y “pirata txirindula” (“bicicleta pirata”, 60 %). En estos cuatro casos, la relación entre los dos elementos de los compuestos no se limita a “un N1-N2 es un N2 que se parece a o que tiene el estilo de un N1”. El compuesto “estrallamar” se podría parafrasear como “una estrella que vive en el mar”, “silla pirata” como “una silla donde se sienta un pirata” y “barco pirata” como “un barco usado por los piratas”. El caso de la “bicicleta pirata” es más ambiguo, ya que la interpretación de “una bicicleta usada por los piratas” no es tan evidente y la de “una bicicleta que tiene el estilo de un pirata” también puede tener cabida. Llegamos a una conclusión similar con el ítem “pirata idazluma” (“bolígrafo pirata”), que obtiene una tasa aún más baja de núcleo a la derecha (28,20 %).¹⁷²

En el caso de otros ítems, es posible que los participantes hayan percibido algunos sustantivos no-nucleares como sustantivos adjetivados que, en este caso, corresponderían también al orden canónico de los sintagmas nominales (N-A) y de los compuestos N-A en euskera (Coyos, 2003, p. 62). En euskera, como en las lenguas romances, ciertos sustantivos también se pueden utilizar como adjetivos y viceversa (Coyos, 2003, p. 49). Como hemos visto en la sección 3.2.2, ciertas palabras son más propicias a la adjetivación en posición de N2 en los compuestos N+N del francés (Arnaud, 2018, 2003). Entonces, podríamos considerar que el mismo proceso de adjetivación existe en euskera, o que nuestros participantes aplican el proceso de adjetivación del francés al euskera. El caso de “pirata” en euskera parece seguir el proceso de “pirata” en español que comentamos en 3.2.1.1. A su vez, “sorpresa” en el ítem “pakete sorpresa” también parece ser un ejemplo claro de este fenómeno, con un 83,3 % de respuestas con el núcleo inicial (inicial).

En cuanto a los dos ítems restantes, la explicación podría ser distinta. Como Liceras et al. (2002, pp. 2-3) mencionan en cuanto a la diferencia entre “letter bomb” en inglés y “carta bomba” en español, el núcleo de un compuesto equivalente se puede identificar de manera diferente según

como “un perro que se parece a un policía”, la imagen de nuestra tarea representaba más bien la segunda opción, por lo que este ítem tenía un significado atributivo como los demás. El octavo ítem animado corresponde a “abeja reina”, que ya hemos comentado anteriormente.

¹⁷² Según este razonamiento, el hecho de que la bicicleta y los bolígrafos no sean asociados muy a menudo a las ocupaciones de los piratas parece ser responsable de la interpretación más bien atributiva de estos compuestos.

la lengua nativa del hablante. En inglés, “*letter bomb*” es considerado como un tipo de bomba, mientras que en español una “carta bomba” es un tipo de carta. La misma lógica se aplica a otro compuesto formado con bomba: “auto bomba” (“*car bomb*” en inglés). Curiosamente, el euskera no parece seguir la lógica del inglés para los compuestos formados con “*bonba*”, sino la del español, como podemos observar con el compuesto atestiguado “*bonba-auto*” (lit., “bomba coche”, registrado en el diccionario *Elhuyar*). Sin embargo, en nuestros datos, encontramos una mayoría de respuestas que siguen el orden opuesto, es decir, un 73,17 % de “*auto-bonba*” y un 55,55 % de “*gutun-bonba*”. En el caso de “*gutun-bonba*”, la tasa se encuentra en torno al 50 %, lo cual se considera generalmente como una posible variación por azar, por lo que nuestros participantes no parecen tener claro cual es el núcleo de ese compuesto.¹⁷³ En cambio, en el caso de “*auto-bonba*”, la tasa es más tajante: o los participantes consideran que el núcleo es “*bonba*”, o consideran que “*bonba*” cumple la función de adjetivo, como hemos mencionada para el compuesto “*pakete sorpresa*”.

Para terminar, si bien no representan la mayoría, las respuestas como (3.126) que incluyen un sustantivo modificado por un participio nos proporcionan una pista de reflexión que implicaría una influencia del francés. En esa lengua, si bien en teoría se podría producir “*voiture-bombe*”, la expresión más común para referirse a un “coche bomba” es “*voiture piégée*” (lit. “coche engañado”), es decir, un sustantivo modificado por un participio, una estructura similar a la encontrada en algunas respuestas en euskera (“*#auto bonbatua*”). Volveremos sobre la direccionalidad de los compuestos en euskera en la discusión, tras el análisis de las respuestas en las demás lenguas del estudio.

En resumen, nuestros participantes produjeron una gran cantidad de compuestos N+N, pero éstos no siempre seguían las normas del euskera en cuanto a la direccionalidad. Además, hemos visto que el rasgo de animacidad afecta el tipo de respuesta y su direccionalidad, ya que los ítems animados aparecen más a menudo como compuestos N+N, pero dicho rasgo también influye en la direccionalidad, al mostrar estos compuestos una direccionalidad opuesta a la norma del euskera. Por su parte, el factor de la(s) lengua(s) hablada(s) en casa tiene un efecto significativo en la direccionalidad de los compuestos, ya que los participantes del grupo FR muestran una preferencia más acentuada por el núcleo inicial, un efecto que podríamos atribuir a la influencia del francés.

¹⁷³ En inglés, se suele usar el término “*at chance level*” para calificar la variación por azar.

Sin embargo, este efecto no se limita al grupo FR, sino que parece afectar, en menor medida, a los participantes de los dos otros grupos. Esta influencia resulta difícil de explicar para los participantes cuya lengua dominante es el euskera, pero no imposible considerando la situación de lengua minorizada del euskera en el País Vasco Francés. Por otra parte, también hemos subrayado la direccionalidad divergente del compuesto “*erle erregina*” que tendremos que tomar en cuenta en nuestros análisis de las demás lenguas.

3.7.1.2 Francés L1/2L1/L2

En francés, de todas las respuestas ($n=832$), encontramos una mayoría de compuestos N+N (60.91 %). Si bien la composición N+N no es considerada como muy productiva en francés, esos resultados no son sorprendentes ya que las imágenes seleccionadas debían favorecer la elicitación de esa estructura y no de compuestos sintagmáticos preposicionales (N+prep+N), una estructura más productiva en esa lengua. Las respuestas que no corresponden a N+N son principalmente estructuras de tipo N+prep+N, como los ejemplos (3.131) (30.23 %), o un sustantivo único acompañado o no de un adjetivo/participio, como los ejemplos (3.132) y (3.133) (7.32 %). Aunque la traducción del ejemplo (3.133) corresponde al orden del compuesto N+N “coche bomba” en español, esta respuesta debe clasificarse como un sustantivo acompañado de un adjetivo ya que contiene un participio usado como un adjetivo y no un segundo N. Las respuestas en blanco y las demás respuestas sin clasificar representan el 1.53 % de los datos de este grupo.¹⁷⁴

(3.131) a.	<i>une chaise de pirate</i>	/	<i>une chaise pirate</i>	
	“una silla de pirata”	/	una silla pirata	
	b.	<i>une étoile de mer</i>	/	<i>une étoile mer</i>
	“una estrella de mar”	/	“una estrellamar”	
(3.132) a.	<i>une bombe</i>	/	<i>une lettre bombe</i>	
	“una bomba”	/	“una carta bomba”	
	b.	<i>un cadeau</i>	/	<i>un paquet surprise</i>
	“un regalo”	/	“un paquete sorpresa”	
(3.133)	<i>une voiture piégée</i>	/	<i>une voiture bombe</i>	
	un coche engañado	/	un coche bomba	
	“un coche bomba”	/	“un coche bomba”	

¹⁷⁴ El caso de “*étoile de mer*” (“estrellamar”) es particular ya que la respuesta esperada en francés era precisamente un N+de+N.

Si miramos más en detalle los ítems en la Figura 3.5, vemos que ciertos ítems favorecen el uso de las preposiciones. Aparte del ítem “*étoile de mer*” (“estrella de mar”), que ha elicitado un 89.66 % de estructuras N+Prep+N, la serie de compuestos que incluyen “pirata” como modificador también tienen una tasa elevada de N+Prep+N: “*bateau pirate*” (“barco pirata”, 69.44 %), “*chaise pirate*” (“silla pirata”, 65.85 %), “*vélo pirate*” (“bicicleta pirata”, 38.46 %) y “*stylo pirate*” (“bolígrafo pirata”, 38.46 %). Son los mismos ítems que correspondían a las tasas más altas de N+Pos+N en euskera, la estructura equivalente de N+Prep+N en francés.

Podemos ver otro patrón que sobresale de la Figura 3.5, la mayoría de los ítems animados se encuentran en la parte superior del gráfico, es decir, tienen una tasa más elevada de compuestos N+N, y los inanimados se ubican en la mitad inferior. Las excepciones son “*paquet surprise*” (“paquete sorpresa”) que tiene una tasa más alta que ciertos animados y “*abeille reine*” (“abeja reina”) que tiene una tasa más baja que algunos inanimados. Para el ítem “*abeille reine*”, podría entrar en juego otro factor como la competición entre la forma N+N y otra forma atestiguada equivalente (“*reine des abeilles*”, “reina de las abejas”).¹⁷⁵ En cuanto al ítem “*paquet surprise*”, como hemos mencionado para el euskera, consideramos que los participantes percibieron el sustantivo “*surprise*” con un alto valor adjetival, lo que favoreció su adjunción a “paquete” sin ningún nexo necesario.

¹⁷⁵ De manera anecdótica, queremos subrayar que el título del cuento de los hermanos Grimm *Die Bienenkönigen* (lit. “La abejas reina”, en inglés “*The queen bee*”) muestra gran variedad en sus traducciones. En español encontramos “La abeja reina” y “La reina de las abejas”, en francés “*La reine des abeilles*”, en euskera “*Erle erregina*” (lit. “abeja reina”) y “*Erleen erregina*” (lit. “abejas-de-las reina”).

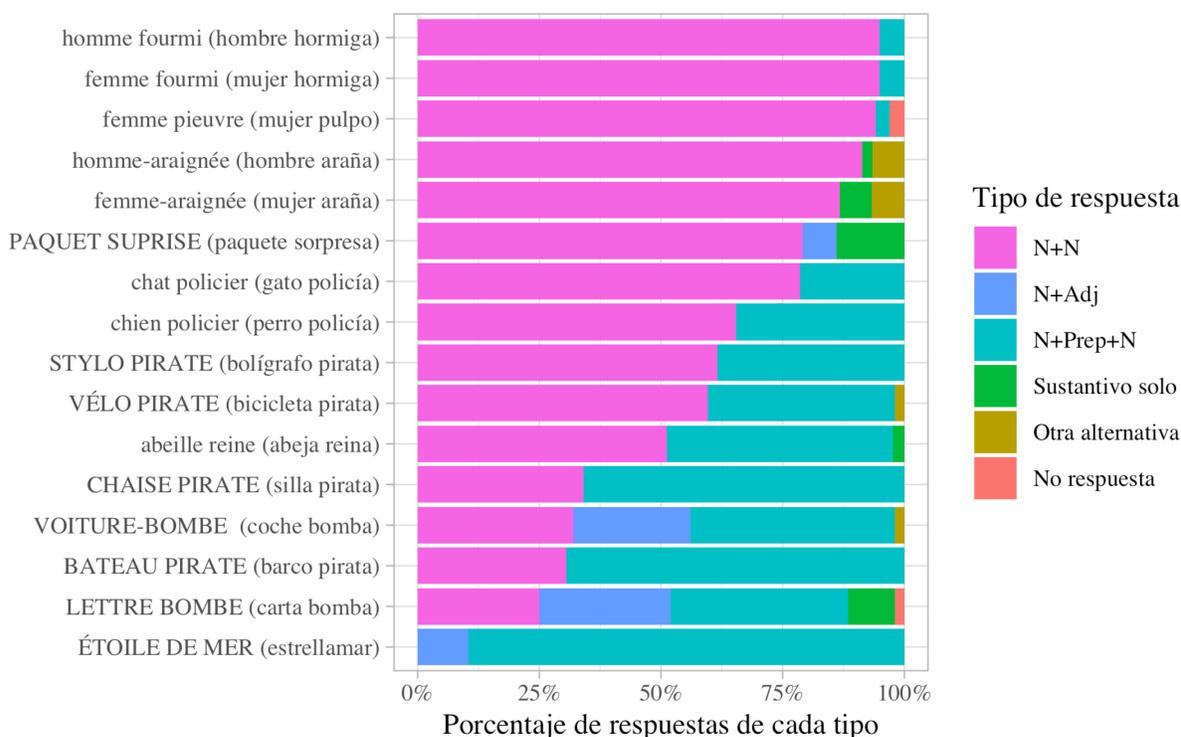


Figura 3.5. Repartición de los tipos de respuestas según los ítems N+N en francés.

Para evaluar los efectos de la animacidad y de los grupos sobre el tipo de respuesta (1=N+N; 0=Otros), creamos un modelo de regresión logística de efectos mixtos siguiendo los mismos pasos que para el euskera. Es decir, incluimos las variables animacidad y grupo como efectos fijos y también la posible interacción entre ellos (animacidad × grupo). En cuanto a la estructura de los efectos aleatorios, incluimos un intercepto aleatorio por participante e ítem y una pendiente aleatoria por participante para la animacidad. Los factores fueron codificados de la siguiente manera (el nivel de referencia corresponde al primer nivel de codificación): animacidad (-.5=inanimado, .5=animado), grupo (1=EU, 2=FR, 3=FR/EU). Dado que ni el factor grupo ni su interacción con el factor animacidad mejoraban la bondad de ajuste del modelo, nuestro modelo final solo incluye el factor fijo de animacidad, que resulta muy significativo ($\beta = 3.2434$, $SE = 0.8820$, $z = 3.677$, $p < .001$). Como en euskera, el valor positivo de este coeficiente nos indica que el rasgo [+animado] favorece la producción de compuestos N+N, lo que apunta, como lo habíamos planteado, a un mayor grado de inseparabilidad de los compuestos animados. A parte del rasgo de animacidad, consideramos la posibilidad de que algunos ítems no fueran propensos a la producción de compuestos N+N ya que no habían sido elegidos precisamente para ese propósito en francés. La

tarea experimental ha sido diseñada con vistas a elicitar compuestos N+N en español y la versión francesa de la tarea consiste en una adaptación con modificaciones míde esta tarea. Si bien nos hemos esforzado en elegir ítems compatibles en las tres lenguas del estudio, es posible que los ítems no sean completamente comparables de una lengua a otra. Volveremos sobre este punto en la discusión (sección 3.8).

En cuanto a la direccionalidad de todas las respuestas que contienen dos sustantivos (N+N, N+Prep+N), encontramos una gran mayoría de respuestas con el núcleo inicial (95.49 %), lo que corresponde al orden canónico del francés. De hecho, casi la totalidad de las respuestas restantes (4,26 %) corresponden a respuestas como “*reine des abeilles*”, que son válidas en francés, ya que, en el sintagma preposicional, el verdadero núcleo es “*reine*” y no “*abeille*”. Para nuestros análisis en español L3, excluirémos este ítem, ya que representaba un caso problemático también en euskera y podría distorsionar nuestros análisis. Excluyendo este ítem, nos quedamos con pocas excepciones con el núcleo final ($n=14$), que están repartidas entre seis ítems, cinco animados y uno inanimado. Reproducimos en (3.134) a (3.138) algunos ejemplos, pero no los comentaremos más ya que se trata de casos marginales. Con una variación tan baja en la direccionalidad de las respuestas, no resulta ni posible, ni relevante analizar los efectos fijos que podrían explicar esa variación.

- (3.134) *une poulpe femme*¹⁷⁶
una pulpo mujer
- (3.135) *une fourmi- femme*
una hormiga mujer
- (3.136) *fourmi homme*
hormiga hombre
- (3.137) *une chipi femme*¹⁷⁷
una chipirón mujer
- (3.138) *une bombe à voiture*
una bomba a/en coche

¹⁷⁶ El sustantivo “poulpe” en francés es masculino, entonces esta respuesta resulta extraña por dos razones: su direccionalidad y el género femenino del artículo “une”.

¹⁷⁷ Si bien en francés la palabra “chipiron” no aparece en el diccionario, es de uso muy común como sinónimo de “calamar/calmar” en la zona del estudio, por un posible contacto con el euskera (“txipiroi”) y el español (“chipirón”). El sustantivo “chipiron” en francés es masculino, entonces esta respuesta con el sustantivo “chipi” que interpretamos como un acortamiento de “chipiron” resulta extraña por las mismas razones que el ejemplo (3.134) explicado en la nota anterior, es decir, su direccionalidad y el género femenino del artículo “une”.

En resumen, los resultados de la prueba de control en francés nos permiten comprobar que, cuando producen compuestos N+N, los participantes los producen con un núcleo inicial, lo que corresponde a lo que se documenta en la literatura sobre el francés. Como hemos notado para la prueba de control en euskera, en cuanto a la productividad, el rasgo [+animado] fomenta la producción de los compuestos N+N.

3.7.1.3 Español L1

Antes de explorar los datos en la tercera lengua de nuestros grupos experimentales, analizamos las producciones de nuestro grupo de control en su primera lengua, el español.

Este análisis nos permite, ante todo, confirmar la validez de la prueba en español, ya que las respuestas elicitadas ($n=336$) corresponden principalmente a compuestos N+N (85.71 %) y, en menor medida, a compuestos sintagmáticos preposicionales (8.16 %), en casi todos los casos con la preposición “de”, como en el ejemplo (3.139). Sin embargo, los ejemplos (3.140) a (3.142) muestran que no solo se usó la preposición “de”, sino también otras, como “con” y “en”. Además, los ejemplos (3.141) a (3.142) nos muestran que con las preposiciones “con” y “en”, nos alejamos de lo que podríamos considerar como un compuesto sintagmático para acercarnos más a lo que son simplemente sintagmas (un sintagma nominal, modificado por un sintagma preposicional). A pesar de ello, para simplificar el análisis de los datos, hemos agrupado todas estas respuestas ya que coinciden en el hecho de presentar una preposición intercalada entre los dos sustantivos.

(3.139) una estrella de mar

(3.140) una carta con bombas

(3.141) explosivos en un sobre

(3.142) un coche azul con bombas

Las demás respuestas se reparten en un 5.78 % de sustantivos solos (modificados o no por un adjetivo) como en los ejemplos (3.143) a (3.145), y un 0.34 % de otras respuestas que, si bien contenían los dos sustantivos, también contenían otro elemento intercalado como lo muestra el ejemplo (3.146).

(3.143) una estrella

(3.144) un bicho humano

(3.145) unos fuegos artificiales

(3.146) un coche azul bomba

Dentro de la categoría de los compuestos N+N, hemos incluido también respuestas que, además del compuesto N+N, contenían otros elementos (por ejemplo, un adjetivo) después del N2 y que no afectaban al significado del compuesto, como se ve en los ejemplos (3.147) y (3.148). Además, hemos aceptado todos los compuestos N+N que llevaban el significado que queríamos elicitarse como lo muestra el ejemplo (3.148), a pesar de la ortografía idiosincrásica que presenta.

(3.147) un barco pirata marrón

(3.148) una personaraña negra (una mujer araña)

Si observamos los 16 ítems de manera individual en la Figura 3.6, podemos ver que el ítem “estrellamar” sobresale con una ausencia total de compuestos N+N, como hemos observado en francés también. Si bien la entrada estrellamar aparece en el diccionario y la hemos incluido dentro de nuestros ítems atestiguados, parece importante reflexionar sobre su pertinencia en la prueba si ninguno de los nativos ha utilizado un compuesto N+N para describirlo. Quitando este ítem, llegamos a un 91.42 % de respuestas que corresponden a un compuesto N+N.

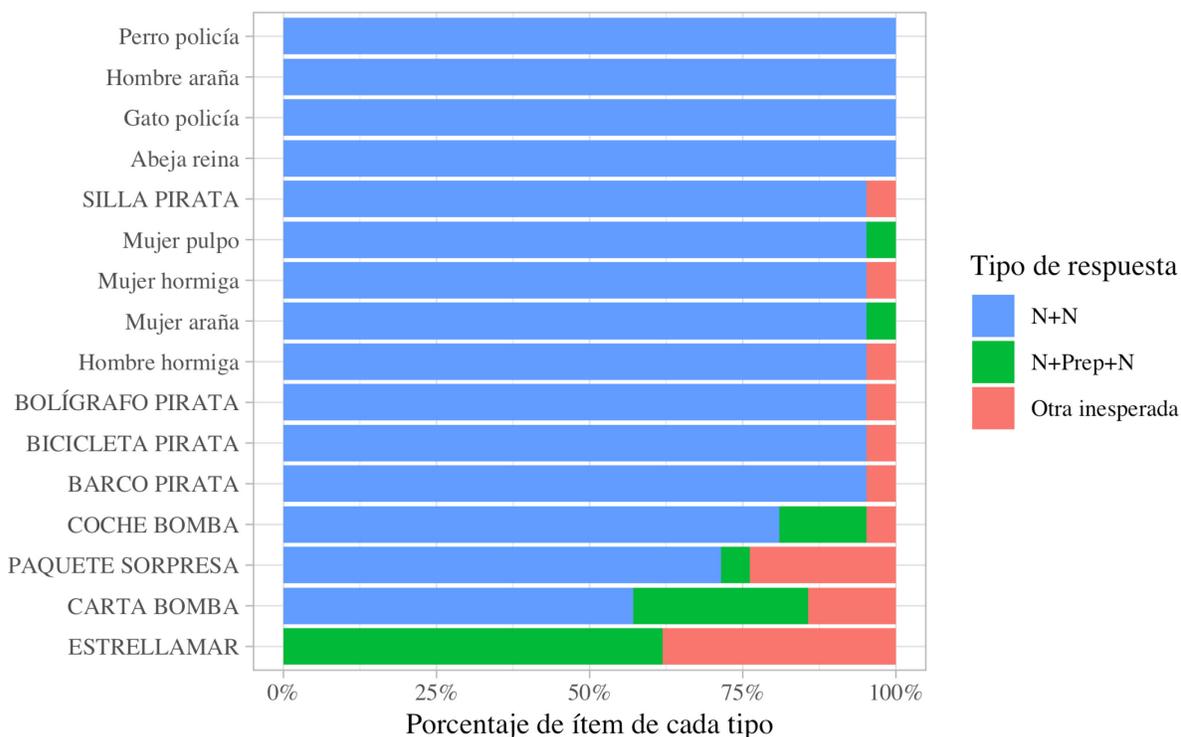


Figura 3.6. Repartición de las respuestas según los ítems N+N en español L1.

A pesar de la poca variación en las respuestas, para evaluar si el efecto de los efectos [+animados] y [+atestiguado] sobre el tipo de respuesta (1=N+N; 0=Otros), aplicamos un modelo de regresión logística de efectos mixtos a nuestros datos siguiendo los mismos pasos que en francés y en euskera, pero con ligeras diferencias en cuanto a los efectos incluidos. En esta ocasión, incluimos las variables animado y atestiguado como efectos fijos y también la posible interacción entre ellos (animado × atestiguado). En cuanto a la estructura de los efectos aleatorios, incluimos un intercepto aleatorio por participante e ítem y una pendiente aleatoria por participante para las variables animado y atestiguado. Los factores fueron codificados de la siguiente manera (el nivel de referencia corresponde al primer nivel de codificación): animacidad (-.5=inanimado, .5=animado), atestiguado (-.5=no atestiguado, .5=atestiguado). Dado que ni el factor atestiguado ni su interacción con el factor animado mejoraban la bondad de ajuste del modelo, nuestro modelo final solo incluye el factor fijo de animacidad, que resulta significativo ($\beta = 6.644$, $SE = 3.340$, $z = 1.989$, $p = .047$). Como en euskera y en francés, el valor positivo de este coeficiente nos indica que el rasgo [+animado] favorece la producción de compuestos N+N, lo que apunta, como habíamos planteado anteriormente, a un mayor grado de inseparabilidad de los compuestos animados. Sin embargo, vemos que, dada la poca variación en las respuestas y la tasa bastante alta de N+N, el efecto es menos significativo que en las dos otras lenguas.

Tras esa breve exploración de los tipos de respuestas presentes en nuestros datos, restringimos nuestro análisis a las respuestas de tipo N+N o N+prep+N –el 95,23 % de los datos– para evaluar la direccionalidad. Como era de esperar, la casi totalidad de las respuestas llevan el núcleo a la izquierda (97.67 %). Las cinco excepciones de N+N aparecen en (3.149) y (3.150), seguidas, después de la barra diagonal, de la respuesta esperada. Las dos excepciones adicionales son como el ejemplo (3.141) y, en realidad, tienen una direccionalidad adecuada ya que son sintagmas nominales complementados por un sintagma preposicional.

- (3.149) a. la reina abeja / una abeja reina
 b. una princesa abispa / una abeja reina
 c. una mosca mujer / una mujer hormiga
 d. un policía gato / un gato policía
 (3.150) una pulpa persona / una mujer pulpo

Cabe destacar aquí que las cinco excepciones aparecen con los compuestos animados y que, si bien no corresponden a lo que la mayoría de los participantes han contestado, las respuestas en

(3.149) se podrían considerar como válidas. Así, es posible que en la mente de esos participantes los estímulos hayan parecido como aparecen en los ejemplos de (3.151).

- | | |
|---------------------------|---|
| (3.151) a. la reina abeja | = ¿la reina con forma de abeja? |
| b. una princesa abispa | = ¿una princesa con forma de avispa de reina? |
| c. una mosca mujer | = ¿una mosca con aspectos de mujer? |
| d. un policía gato | = ¿un policía con aspectos de gato? |

Constatamos aquí que la frontera es borrosa cuando llegamos a la interpretación de lo que es el núcleo de un compuesto N+N animado, incluso para los nativos de español. En cambio, si bien podríamos seguir la misma lógica para el ejemplo (3.150), no encontramos ninguna justificación para la flexión de género en la palabra pulpo, un sustantivo epiceno que debería conocer un hablante nativo. Fuera del análisis descriptivo de esas excepciones en los datos, resulta imposible analizar la casi ausencia de variación en la direccionalidad con un modelo estadístico.

En resumen, los datos de nuestro grupo de control nos revelan que la prueba permite elicitar principalmente las respuestas esperadas, es decir, compuestos de tipo N+N con un núcleo inicial. El análisis individual de los ítems nos ha permitido concluir que deberíamos descartar el ítem “estrellamar” porque en los hablantes nativos elicitaba respuestas como “estrella” o “estrella de mar”, pero en ningún caso el compuesto N+N “estrellamar”.

3.7.1.4 Español L3: las etapas iniciales

Antes de continuar con los datos del español como L3, resulta necesario ajustar nuestras hipótesis conforme a lo que hemos descubierto sobre los compuestos N+N en las dos primeras lenguas de nuestros participantes y hacer algunos ajustes a nuestro juego de datos.

En cuanto al juego de datos, hemos visto que el ítem “reina abeja” era problemático en francés y en euskera, y que, a su vez, el ítem “estrellamar” presentaba problemas en francés y en euskera. Por lo tanto, descartamos estos dos ítems de nuestro análisis del español L3. A modo recapitulativo, presentamos en la Tabla 3.9 las tasas de productividad de los N+N y de direccionalidad sin estos dos ítems en las lenguas nativas o casi nativas de los participantes.

Tabla 3.9. Productividad y direccionalidad de los N+N en las lenguas (casi) nativas de los participantes.

	Euskera	Francés	Español
Producción de N+N	80,08 %	64,67 %	90,82 %
Direccionalidad canónica	39,47 %	97,35 %	98,20 %

En cuanto a la productividad de los compuestos N+N, vemos que la prueba favoreció una producción mayoritaria de N+N en las tres lenguas. Sin embargo, la tasa del francés es más baja, algo que podríamos atribuir al diseño de la prueba de control en francés y, en menor medida, en euskera. Efectivamente, si bien nos hemos esforzado para diseñar una prueba compatible con las tres lenguas del estudio, también es cierto que hemos otorgado más importancia a los ítems en la lengua meta de los participantes de L3, el español. Por ejemplo, hemos conservado ítems como “coche bomba” a pesar de que el equivalente más común en francés no correspondía a un compuesto N+N. El razonamiento seguido en este caso era que, en principio, nada en francés impide decir “*voiture bombe*”, en lugar del más común “*voiture piégée*”.

Ahora, volviendo a la hipótesis N1 y considerando las tasas relativamente altas de productividad de los N+N en las dos primeras lenguas de nuestros participantes, es posible que no encontremos una influencia aparente en lo que respecta a la tasa de producción de N+N en español como L3, sea cual sea la lengua de influencia. Sin embargo, la hipótesis inicial se mantiene, ya que la tasa de N+N en euskera es más alta que la del francés.

Por lo que concierne a la direccionalidad (hipótesis N2), hemos visto que, incluso en euskera, los participantes del grupo FR producían más N+N con el núcleo inicial, lo que no corresponde al orden canónico de los compuestos N+N en euskera. Es decir, la influencia del francés parece aplicarse ya a la L2 de esos participantes, por lo que es más probable que se extienda a la L3 también.

En lo relativo al rasgo de animacidad, hemos visto que, en las tres lenguas, los ítems animados favorecen la formación de compuestos N+N, por lo que podemos esperar una mayor tasa de N+N con los ítems animados, como planteábamos en nuestra hipótesis N3. Sin embargo, dado que las tres lenguas muestran los mismos patrones, resultará difícil aclarar si se trata de una tendencia universal (o al menos una tendencia en las tres lenguas del estudio) y no podremos

atribuir la influencia a una lengua en concreto. Hemos visto también que la animacidad favorecía la producción de respuestas con núcleo final en euskera. En cambio, la animacidad no tenía ningún efecto significativo sobre la direccionalidad de las respuestas en francés o en español.

Con respecto a la hipótesis N4 sobre el hecho que un compuesto esté atestiguado o no, no hemos encontrado ningún efecto en español L1, dado que había poca variación en los datos. En cuanto al euskera y el francés, no era un factor que entraba en nuestro diseño. Entonces, mantenemos la hipótesis N4 sin modificaciones y seguimos considerando que, si se observa un efecto, los ítems atestiguados deberían elicitar una mayor tasa de respuestas esperadas (N+N con núcleo inicial) para todos los participantes.

Una vez ajustadas las hipótesis, pasamos al análisis de nuestros grupos experimentales de español L3. Los resultados muestran un patrón semejante al del grupo de control, con una mayoría de respuestas que incluían compuestos N+N (75,27 %) y, en menor medida, N+Prep+N (9,07 %), como muestra la Figura 3.7. La diferencia en la tasa de producción de compuestos N+N entre el grupo de control y los grupos experimentales asciende al 15,54 %. En cuanto a la direccionalidad, la diferencia es menor, es decir, llega solo a un 9,3 %.

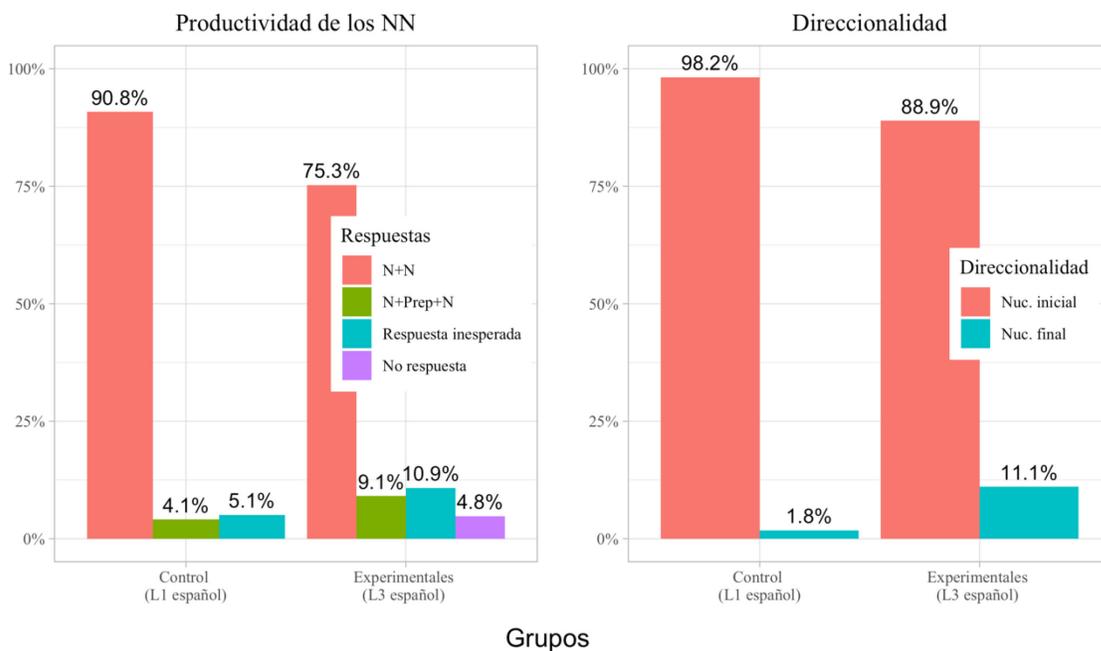


Figura 3.7. Distribución de las respuestas para los ítems N+N. Todos los grupos experimentales (EU, FR, FR/EU) aparecen juntos.

Como en los análisis anteriores (euskera, francés y español nativos y casi nativos), un primer escrutinio visual de los 14 ítems de manera individual en la Figura 3.8 nos permite ver que los ítems animados tienen una tasa más alta de compuestos N+N.

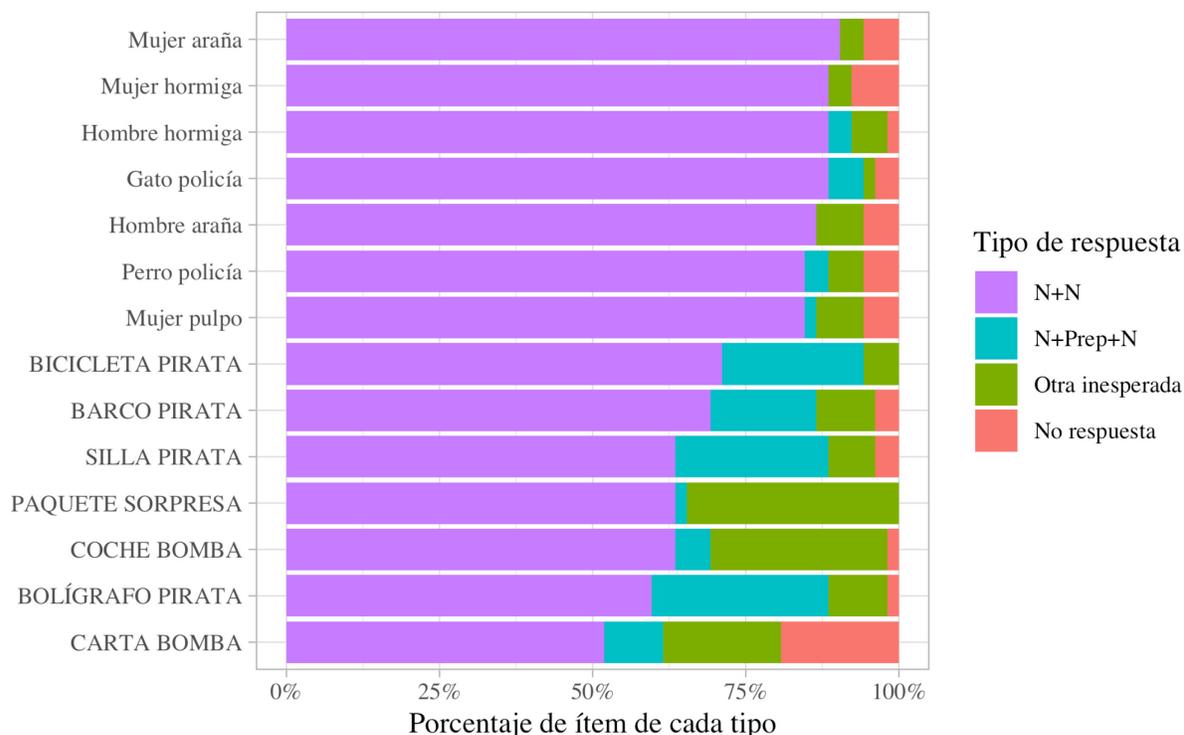


Figura 3.8. Repartición del tipo de respuesta según el ítem en español L3.

Cabe destacar también que, si bien los únicos compuestos sintagmáticos preposicionales aparecían con cinco ítems en el grupo de control, ahora ocurren con casi todos los ítems, aunque de manera variable. Los tres ítems que incluyen “*pirate*” (“pirata”) como segundo elemento son los que muestran la mayor tasa de compuestos sintagmáticos preposicionales, como ocurría también en francés.¹⁷⁸ Como en el análisis de los datos de español L1, dentro de la categoría de los compuestos N+N, hemos incluido también respuestas que, además del compuesto N+N, contenían otros elementos (por ejemplo, un adjetivo) después del N2 y que no afectaban al significado del compuesto, como muestran los ejemplos (3.152) y (3.153). Además, no hemos tenido en cuenta los errores ortográficos, ni los guiones, ni los espacios.

¹⁷⁸ El sustantivo “*pirate*” parece seguir un camino diferente de sus equivalentes “pirata” en español y “*pirata*” en euskera que muestran una fuerte tendencia a la adjetivación.

- (3.152) señora poulpa violeta
 (3.153) un paquete sorpresa marron

Para evaluar de manera conjunta los efectos de animacidad, grupo y atestiguado sobre el tipo de respuesta (1=NN; 0=Otros), elaboramos un modelo de regresión logística de efectos mixtos, en el que incluimos todas estas variables como efectos fijos y también todas las interacciones posibles entre esos tres efectos. En cuanto a la estructura de los efectos aleatorios, tuvimos que reducirla a un intercepto aleatorio por participante e ítem y una pendiente aleatoria por participante para la animacidad para que el modelo convergiera. Los factores fueron codificados de la siguiente manera (el nivel de referencia corresponde al primer nivel de codificación): animacidad (-.5=inanimado, .5=animado), atestiguabilidad (-.5=inventado, .5=atestiguado), grupo (1=ES, 2=EU, 3=FR, 4=FR/EU). Los resultados del modelo, que no incluye ninguna interacción ni el efecto atestiguado porque no mejoraban la bondad de ajuste del modelo, aparecen en la Tabla 3.10.

Tabla 3.10. Resultados del modelo regresión logística de efectos mixtos del tipo de respuesta para los estímulos N+N.

	β	<i>SE</i>	<i>z</i>	<i>Pr(> z)</i>
(Intercepto)	3.945	0.593	6.658	<.001
Animado	3.258	0.704	4.631	<.001
Grupo (EU)	-0.982	0.746	-1.317	.188
Grupo (FR)	-1.450	0.734	-1.976	.048
Grupo (FR/EU)	-1.644	0.652	-2.522	.012

Este modelo nos confirma que el grupo de control produce significativamente más compuestos N+N que dos de los grupos experimentales (FR y FR/EU). Una prueba *posthoc* de comparaciones múltiples de medias (Contraste Tukey) nos revela que, a pesar de la ausencia de diferencia significativa del grupo EU con respecto al grupo de control, los grupos experimentales no muestran ninguna diferencia significativa entre sí en cuanto a la tasa de producción de N+N (FR – EU: $\beta = -0.468$, $SE = 0.783$, $z = -0.598$, $p = .55$; FR/EU – EU: $\beta = -0.662$, $SE = 0.711$, $z = -0.931$, $p = 0.35$; FR/EU – FR: $\beta = -0.194$, $SE = 0.696$, $z = -0.280$, $p = 0.78$). La diferencia entre el grupo de control y los grupos experimentales se explica parcialmente por una tasa de omisiones (4,80 %) en los grupos de L3 que no encontramos en el grupo de control y por una tasa más elevada de sustantivos solos, modificados o no por un adjetivo (9,20 %). El porcentaje de respuestas de otro tipo, como por ejemplo dos sustantivos con otro elemento intercalado (“*un coche azul bomba”),

es marginal tanto en el grupo de control como en los grupos experimentales (1.65 %). En cuanto al factor animacidad, como en nuestros análisis en otras lenguas, la Tabla 3.10 muestra que la condición [+animado] fomenta la producción de compuestos N+N de manera significativa. Esa condición afecta a todos los grupos de manera similar, ya que el hecho de incluir la interacción grupo \times animacidad no mejoraba la bondad de ajuste del modelo.

Seguimos los mismos pasos para evaluar los efectos de grupo, animacidad y atestiguado sobre la direccionalidad de los compuestos N+N, pero conservando solamente las respuestas correspondientes a la estructura N+N. El modelo resultante contiene los mismos efectos fijos y aleatorios. La diferencia principal se encuentra en la amplitud y la significatividad de los efectos fijos. El efecto animacidad ahora es negativo ($\beta = -4.665$, $SE = 2.720$, $z = -1.715$, $p = 0.08$), pero solo marginalmente significativo, es decir, la condición [+animado] fomenta marginalmente la producción de compuestos con un núcleo a la derecha. El efecto de grupo es significativo para los grupos EU ($\beta = -3.329$, $SE = 1.440$, $z = -2.312$, $p = 0.02076$) y FR/EU ($\beta = -3.954$, $SE = 1.325$, $z = -2.984$, $p = 0.00285$), pero no lo es para el grupo FR ($\beta = -2.332$, $SE = 1.487$, $z = -1.569$, $p = 0.11664$). De manera general, los grupos experimentales producen menos compuestos con la direccionalidad correcta que el grupo de control, pero el grupo FR es el que se acerca más al grupo de control.

En resumen, el análisis de esos datos nos permite ver que los grupos experimentales se diferencian del grupo de control, tanto en cuanto a la productividad de los compuestos N+N como a la direccionalidad de sus respuestas. El grupo EU es el que más se acerca del grupo de control para la productividad de los N+N. En cambio, el grupo FR es el más parecido al grupo de control en cuanto a la direccionalidad. El rasgo [+animado] fomenta la producción de compuestos N+N, efecto que observamos en todas las lenguas del estudio. Además, este efecto coincide con el mayor grado de fusión de los N+N animados observado por Liceras y Klassen (2019), como mencionamos en la sección 3.4.3. Curiosamente, el mismo rasgo [+animado] parece dificultar la identificación del núcleo, ya que tiene un impacto negativo sobre la direccionalidad de los compuestos, algo que también hemos observado en los resultados en francés y en euskera de nuestros participantes. Los compuestos animados parecen tener una direccionalidad casi reversible a nivel de interpretación ya que sus dos elementos tienen un mayor grado de fusión, algo que no observamos con los inanimados. A su vez, el hecho de que un compuesto esté atestiguado o no no tiene ningún efecto significativo sobre la producción de nuestros participantes, ni en el tipo de respuesta, ni en la

direccionalidad. Esta ausencia de efecto no es sorprendente, ya que los participantes de los grupos experimentales habían recibido una exposición limitada al español. Sin embargo, con más horas de instrucción, es posible que observemos un efecto en las recolecciones posteriores.

3.7.1.5 Español L3: perspectiva longitudinal

Para ver cómo evoluciona el dominio del fenómeno con el paso del tiempo en nuestros grupos experimentales, comparamos los datos de la primera recolección con las dos recolecciones sucesivas que se llevaron a cabo después de 140 y 200 horas de enseñanza en español, respectivamente.

Para comprobar el efecto de cada variable presente en nuestro análisis hasta ahora, construimos un modelo para cada recolección de datos con todos los factores relevantes siguiendo los pasos de los modelos de la primera recolección, con la misma prueba *posthoc* para determinar las diferencias entre los grupos experimentales. Además, construimos un modelo global para las tres recolecciones incluyendo el factor etapa de recolección, para comprobar la evolución en el tiempo.

Como vemos en la Figura 3.9, la tasa de producción de N+N (parte izquierda) no aumenta de manera significativa entre la primera y la segunda etapa de recolección ($\beta = 0,155$, $SE = 0.149$, $z = 1.038$, $p = .299$), pero sí lo hace entre la segunda y la tercera ($\beta = 1.272$, $SE = 0.179$, $z = 7.097$, $p < .001$). Por su parte, la tasa de respuestas con el núcleo inicial (parte derecha de la figura) aumenta de manera significativa a partir de la segunda recolección de datos ($\beta = 3.5023$, $SE = 0.501$, $z = 6.998$, $p < .001$), pero la diferencia entre la segunda y la tercera recolección no es significativa ($\beta = -0.023$, $SE = 0.547$, $z = -0.042$, $p = .967$) porque los tres grupos casi alcanzan el 100 % a partir de la segunda recolección.

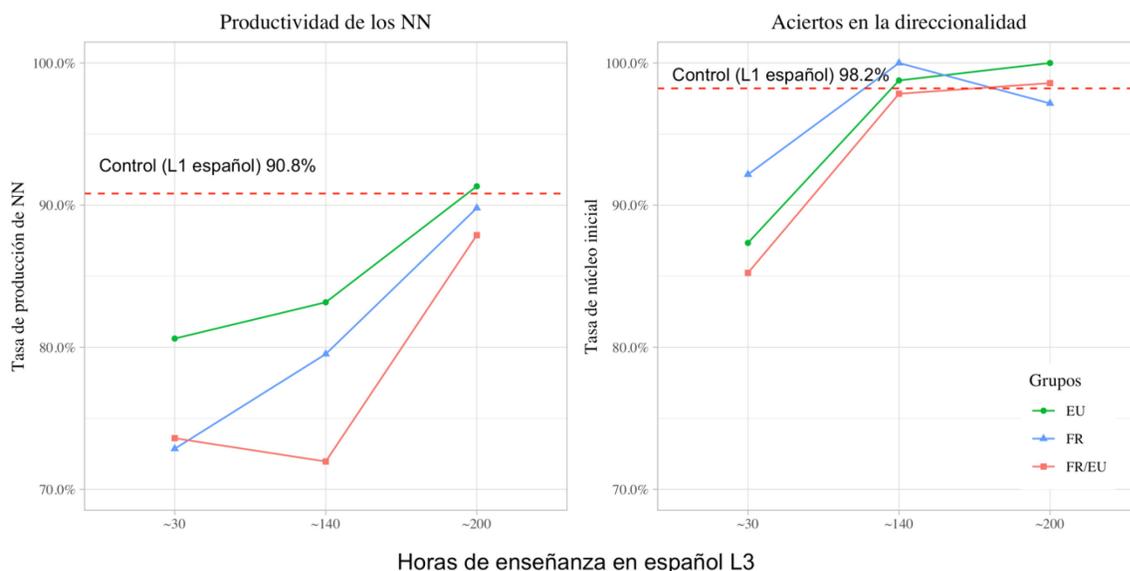


Figura 3.9. Evolución de la productividad y la direccionalidad de los NN - Grupos experimentales (la escala en Y comienza a partir de 70 % para hacer más visibles las diferencias entre los grupos).

En lo referente a la evolución de la productividad etapa por etapa, los grupos experimentales muestran una evolución divergente en la segunda recolección. Concretamente, el grupo FR/EU muestra una diferencia significativa con el grupo de control ($\beta = -1.642$, $SE = 0.804$, $z = -2.043$, $p < 0.041$), pero la diferencia entre este y los grupos experimentales deja de ser significativa a partir de la segunda recolección. Los tres grupos experimentales muestran una evolución comparable en cuanto a la direccionalidad, y, de nuevo, las diferencias entre los grupos desaparecen a partir de la segunda recolección. Con respecto al factor animacidad, su efecto se mantiene de manera significativa solo sobre la productividad. En la segunda recolección sigue siendo muy significativo ($\beta = 3.885$, $SE = 0.996$, $z = 3.901$, $p < .001$) y apenas disminuye en la tercera etapa ($\beta = 5.199$, $SE = 1.798$, $z = 2.892$, $p = 0.004$).

Resumiendo, no observamos grandes dificultades en cuanto a la direccionalidad de manera general, lo cual parece indicar un efecto facilitador del francés, que muestra características semejantes al español. Las pequeñas diferencias observadas en cuanto al rasgo de animacidad persisten en el tiempo, pero considerando lo que hemos observado en euskera, en francés y en español, resulta difícil relacionar este fenómeno con una influencia de una lengua en concreto. De hecho, parece observarse un efecto de interpretación que se aplica a las tres lenguas de nuestro estudio, concretamente, los compuestos N+N animados muestran un mayor grado de fusión. A

continuación, presentamos el análisis de los compuestos deverbales, siguiendo las mismas etapas que en la presente subsección. Tras ello, pasaremos a la discusión general de los resultados.

3.7.2 Los deverbales

3.7.2.1 Euskera L1/2L1/L2

Las respuestas en euskera ($n=832$) corresponden principalmente a compuestos deverbales (76,80 %), pero encontramos también un 22,36 % de producciones que corresponden a formas inesperadas y un 0,84 % sin respuesta.

Dentro de las formas inesperadas, los compuestos formados a partir de un sustantivo y una forma verbal de base son poco numerosos (0,96 %), pero de manera general respetan el orden de los componentes de los compuestos deverbales del euskera, es decir, N+V. Por ejemplo, encontramos formas como *lursagar zuritu* o *bixkotx aurtiki*, que corresponderían respetivamente a “pelar patata” y “lanzar pastel”.¹⁷⁹ En cuanto a las otras formas que no corresponden a un compuesto del tipo Objeto-Verbo o Verbo-Objeto, encontramos una variedad de respuestas. Varias respuestas consisten en un sustantivo solo (10,94 %) y conllevan buena parte del significado del compuesto que queríamos elicitarse, como muestran los ejemplos de (3.154). Otras respuestas corresponden simplemente al deverbale solo (6,37 %), como muestran los ejemplos de (3.155). Otras respuestas adoptan la forma de un compuesto N+N (0,96 %), como el ejemplo (3.156). Finalmente, el resto (3,13 %) corresponde o a construcciones incompletas o a estructuras más complejas (como oraciones completas). En los ejemplos (3.154) a (3.156), la expresión después de la barra diagonal corresponde al compuesto deverbale esperado en euskera.

- (3.154) a. *katapulta* / *bixkotx aurtikigailu*
 catapulta / *pastel lanzador*
 “lanzapasteles”
- b. *marteilu* / *botila hautsigailu*
 martillo / *botella rompedor*
 “rompebotella”
- c. *dentista* / *hortz ateratzailea*
 dentista / *diente arrancador*
 “arrancadientes”
- (3.155) a. *garbitzaile* / *txirindula garbitzaile*

¹⁷⁹ Nuestra interpretación es aproximada, ya que son formas idiosincrásicas.

	lavador (ag.)	/	bicicleta	lavador
	“lavabicycletas”			
b.	<i>garbigailu</i>	/	<i>azieta</i>	<i>garbigailu</i>
	lavador (instr.)	/	plato	lavador
	“lavaplatos”			
c.	<i>zaintzaile</i>	/	<i>oihan</i>	<i>zain(tzailea)</i>
	guardador	/	bosque	guardador
	“guardabosques”			
(3.156)	<i>aiztur</i>		<i>kontserba</i>	/
	tijera		lata	/
			lata	abridor
	“abrelatas”			

Cabe destacar aquí que, si bien conservamos la oposición animado-agentivo vs. inanimado-instrumental en los ejemplos (3.155) (-*gailu* para un inanimado, -*tzaile* para un animado), siguiendo a Azkarate (1995), los participantes no siempre hicieron esta distinción, como veremos más en detalle a continuación.

Volviendo al contraste entre los compuestos deverbales y los demás tipos de respuestas, se puede observar cierta variación en función de los ítems. La Figura 3.10 permite ver que los ítems inanimados (en mayúscula) tienen una menor tasa de compuestos deverbales, mientras que los ítems animados tienen una tasa un poco más alta. La principal excepción a esta tendencia es el ítem “*hortz arrapatzaile*” que ha sido reemplazado por el sustantivo “*dentista*” por una mayoría de participantes.

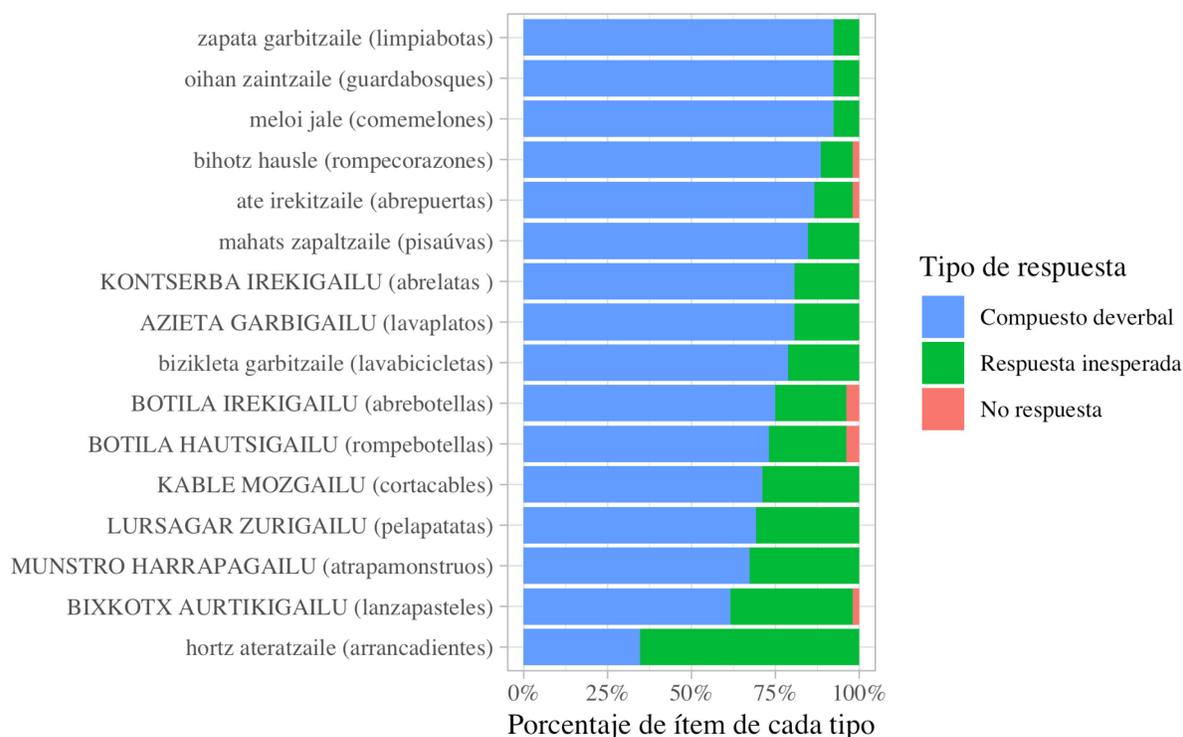


Figura 3.10. Repartición de las respuestas según los ítems de compuestos de verbales en euskera.

De manera global, el promedio de compuestos de verbales es de un 81,25 % en los ítems animados y baja a un 72,36 % en los inanimados. Para evaluar los efectos de animacidad y de grupos sobre el tipo de respuesta (1=Compuesto de verbal; 0=Otros), elaboramos un modelo de regresión logística de efectos mixtos siguiendo los mismos pasos que en los análisis anteriores. Es decir, incluimos las variables animacidad y grupo como efectos fijos y también la posible interacción entre ellos (animacidad \times grupo). En cuanto a la estructura de los efectos aleatorios, incluimos un intercepto aleatorio por participante e ítem y una pendiente aleatoria por participante para la animacidad. Los factores fueron codificados de la siguiente manera (el nivel de referencia corresponde al primer nivel de codificación): animacidad (-.5=inanimado, .5=animado), grupo (1=EU, 2=FR, 3=FR/EU). Si bien el modelo completo convergía, ninguno de los efectos fijos era significativo. Además, la eliminación de cada efecto fijo no afectaba de manera considerable la bondad de ajuste del modelo, por lo que concluimos que ninguno de los efectos fijos considerado era significativo.

Si bien no se observa ningún efecto de grupo o de animacidad sobre los tipos de respuestas producidas a nivel global, seguimos indagando en la morfología de los deverbales para contrastar lo que dice la literatura con respecto a los sufijos deverbales.

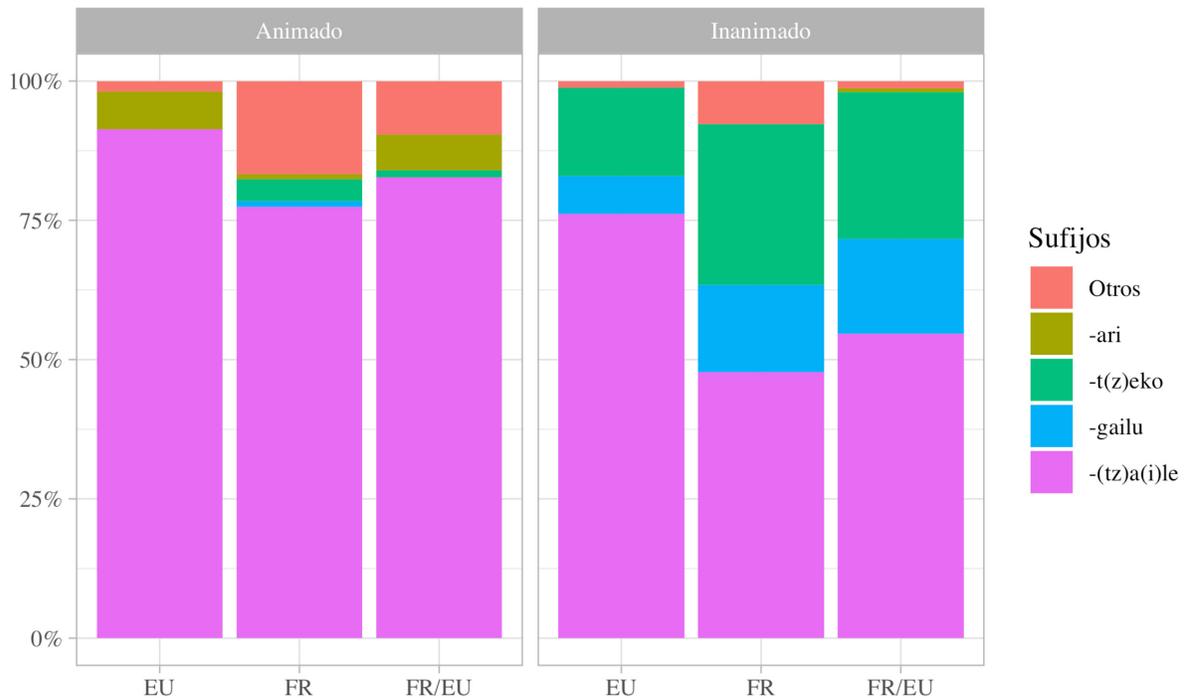


Figura 3.11. Repartición de las respuestas según la morfología de los deverbales, por grupo y por animacidad en euskera.

La Figura 3.11 muestra los sufijos utilizados en los compuestos deverbales y en los deverbales solos ($n=692$), según el rasgo de animacidad de los ítems y los grupos de participantes. El sufijo *-tzaile* (o sus alomorfos) es el más frecuente, tanto para los animados como los inanimados. El sufijo *-ari* aparece solo en los deverbales animados. Por su parte, el sufijo instrumental *-gailu* se usa menos de lo esperado y, además del sufijo *-tzaile*, tiene otro competidor en el sufijo *-t(z)eko*.

Si bien hemos clasificado las respuestas con el sufijo *-tzeko* dentro de los compuestos deverbales, cabe aclarar que, a diferencia de los compuestos deverbales que llevan un objeto escueto (sin ninguna posposición), los adnominales formados con el sufijo *-tzeko* pueden llevar algunos elementos intercalados entre el objeto y el verbal. En nuestros datos, algunas respuestas en *-tzeko* llevan un objeto con un determinante afijado (singular o plural), lo que no era el caso con

los demás deverbales formados a partir de los sufijos *-tzaile*, *-gailu* y *-ari*, como ilustramos en (3.157).

- | | | |
|------------|------------------|----------------------|
| (3.157) a. | <i>munstroak</i> | <i>harrapatzeko</i> |
| | monstruo.DET.PL | atrapador |
| b. | <i>munstro</i> | <i>harrapagailu</i> |
| | monstruo | atrapador |
| c. | <i>munstro</i> | <i>harrapatzaile</i> |
| | monstruo | atrapador |

Con el fin de comprobar las variables que influyen sobre el uso de cada sufijo, creamos un modelo de regresión logística de efectos mixtos con cada tipo de sufijo, siguiendo los mismos pasos que en el análisis anterior. En el primer modelo con el sufijo *-tzaile* (*-tzaile*=1, otros=0), dado que la interacción animado \times grupo no mejoraba la bondad de ajuste del modelo, la eliminamos. En el modelo resultante, el efecto animado es significativo y positivo ($\beta = 1.509$, $SE = 0.7601$, $z = 1.985$, $p = .047$), es decir que, se usa significativamente más el sufijo *-tzaile* con los ítems animados. En cuanto al efecto de grupo, una prueba *posthoc* de comparaciones múltiples de medias (contraste *Tukey*) nos revela que el contraste FR – EU es significativo ($\beta = -1.475$, $SE = 0.59$, $z = -2.498$, $p = .033$), mientras que los contrastes FR/EU – EU ($\beta = -1.0573$, $SE = 0.545$, $z = -1.942$, $p = .126$) y FR/EU – FR ($\beta = 0.4172$, $SE = 0.504$, $z = 0.828$, $p = .684$) no lo son.

Siguiendo los mismos pasos, comprobamos los sufijos *-gailu*, *-tzeko* con un modelo para cada uno, pero como los modelos no convergían, tuvimos que eliminar la pendiente aleatoria por participante para la animacidad. A partir de los modelos resultantes, como el efecto de grupo no mejoraba la bondad de ajuste del modelo lo descartamos, por lo que reportamos aquí solo el efecto de animacidad para cada modelo. El efecto del factor animacidad es negativo y significativo para los sufijos *-gailu* ($\beta = -4.962$, $SE = 1.148$, $z = -4.324$, $p < .001$) y *-tzeko* ($\beta = -3.942$, $SE = 0.646$, $z = -6.099$, $p < .001$), es decir que, estos sufijos se usan significativamente menos para los ítems animados.

El modelo para el sufijo *-ari* se parece más al modelo de *-tzaile*, en el sentido que incluye el factor grupo y que el efecto del factor animacidad es positivo. Así, el sufijo *-ari* aparece significativamente más en los ítems animados ($\beta = 2.810$, $SE = 1.168$, $z = 2.406$, $p = .016$) y el efecto del grupo FR es marginalmente significativo ($\beta = -2.191$, $SE = 1.236$, $z = -1.772$, $p = .076$), es decir que, el grupo FR usa menos el sufijo *-ari* que el grupo EU. Una prueba *posthoc* de

comparaciones múltiples de medias (contraste *Tukey*) nos revela, sin embargo, que los demás contrastes entre los grupos no son significativos.

Para completar el análisis de los sufijos, construimos un último modelo siguiendo los mismos pasos para evaluar los efectos de animacidad y grupo sobre el uso de los demás sufijos que habíamos clasificado en la categoría “Otros” (Otros=1, *-ari*, *-tzeko*, *-gailu*, *-tzaile* = 0). Este modelo nos revela que el efecto de animacidad es positivo y significativo ($\beta = 5.925$, $SE = 2.279$, $z = 2.600$, $p = .009$), lo que quiere decir que se usan más sufijos idiosincrásicos con los ítems animados. El efecto del grupo FR es también positivo y significativo ($\beta = 2.919$, $SE = 1.208$, $z = 2.417$, $p = 0.016$), es decir que, el grupo FR usa más estos sufijos idiosincrásicos que el grupo EU. Una prueba *posthoc* de comparaciones múltiples de medias (contraste *Tukey*) nos revela que los demás contrastes entre los grupos no son significativos.

Resumiendo, observamos dos patrones de especial interés en el uso de los sufijos deverbales en las repuestas de nuestros participantes. Por una parte, los sufijos *-tzaile* y *-ari* se utilizan más con los ítems animados, mientras que los sufijos *-tzeko* y *-gailu* son más frecuentes con los ítems inanimados. La oposición *-tzaile/-gailu* corresponde a lo que se documenta en la literatura revisada en la sección 3.2.3.2. Sin embargo, el sufijo *-tzaile* también es usado por nuestros participantes para referirse a compuestos inanimados, algo que está menos documentado, y el sufijo *-tzeko* aparece ocasionalmente en usos poco documentados. Por otra parte, notamos una diferencia en el uso de los sufijos por parte de los participantes del grupo FR. Éstos usan significativamente menos los sufijos *-tzaile* y *-tzeko*, pero producen significativamente más formas idiosincrásicas de deverbales que los dos otros grupos. A modo de ejemplo, reproducimos en (3.158) dos de esas formas idiosincrásicas que resultan difíciles de interpretar.

- (3.158) a. *melo* *jaten*
 melón comer.PRS
 b. *ate* *irekitzea*
 puerta apertura

En resumen, nuestros participantes produjeron una gran cantidad de compuestos deverbales que, de manera general, seguían las normas del euskera, tanto en cuanto a la direccionalidad (N+V+sufijo) como en cuanto a la sufijación del verbal (*-tzaile*). Sin embargo, aparecieron diferencias entre los grupos en cuanto a los sufijos utilizados para formar los deverbales y el grupo FR se destacó por su uso de algunas formas idiosincrásicas difíciles de analizar. Además, apareció

en nuestros datos el sufijo adnominal *-tzeke*, menos frecuente en la literatura sobre los deverbales, cuestión sobre la que volveremos en la discusión. Cabe recordar que, en menor medida, también se produjeron otras formas no esperadas, como un sustantivo o un sustantivo deverbale solo como en los ejemplos de (3.155), lo cual correspondía también a una opción válida en la mayoría de los casos, siempre cuando conllevaba un significado similar al del compuesto deverbale esperado.

3.7.2.2 Francés L1/2L1/L2

Las respuestas en francés corresponden principalmente a compuestos sintagmáticos preposicionales que contienen un verbo derivado (47,12 %; de aquí en adelante, compuestos preposicionales) y compuestos deverbales propios (28,25 %), pero también tenemos un alto porcentaje de producciones que corresponden a otras formas inesperadas (23,44 %) y un 1,20 % sin respuesta.

En cuanto a las formas alternativas que no corresponden a un compuesto del tipo Objeto-Verbo o Verbo-Objeto, encontramos una variedad de respuestas. Como en euskera, varias respuestas consisten en un sustantivo solo (12,01 %) que conlleva buena parte del significado del compuesto que queríamos elicitare, como muestran los ejemplos de (3.159). Otras respuestas corresponden simplemente a un verbo derivado mediante un sufijo (5,64 %), como en (3.160). Otras respuestas (1,20 %) adoptan la forma de compuestos N+N o N-à/de-N, como en (3.161). Finalmente, el resto (4,57 %) corresponde o a construcciones incompletas o a estructuras más complejas (como oraciones completas).

(3.159) a.	<i>piège</i>	/ attrape-	<i>monstre</i>
	trampa	/ atrapa-	monstruo
	“atrapamonstruos”		
b.	<i>marteau</i>	/ casse-	<i>bouteille</i>
	martillo	/ rompe-	botella
	“rompebotellas”		
c.	<i>dentiste</i>	/ arrache-	<i>dent</i>
	dentista	/ arranca-	diente
	“arrancadientes”		
(3.160) a.	<i>éplucheur</i>	/ épluche-	<i>patate</i>
	pelador	/ pela-	patata
	“pelapatatas”		
b.	<i>ouvre</i>	/ ouvre-	<i>porte</i>
	abridor	/ abre-	puerta
	“abrepuertas”		

- (3.161) a. *homme porte* / *ouvre- porte*
 hombre puerta / abre- puerta
 “abrepuertas”
- b. *piège à/de monstre* / *attrape- monstre*
 trampa para monstruos / atrapa- monstruo
 “atrapamonstruos”

Tras esa breve exploración de las respuestas inesperadas, que representan un 23,44 % de los datos, volvemos a los tipos de respuesta más frecuentes. En la Figura 3.12, se puede notar que ciertos ítems favorecen la producción de compuestos deverbales, mientras que otros parecen favorecer los compuestos preposicionales. De entrada, el hecho de que encontremos más compuestos preposicionales (47,12 %) que compuestos deverbales propios (28,25 %) es sorprendente, ya que, según la literatura, esta última estructura se considera muy productiva.

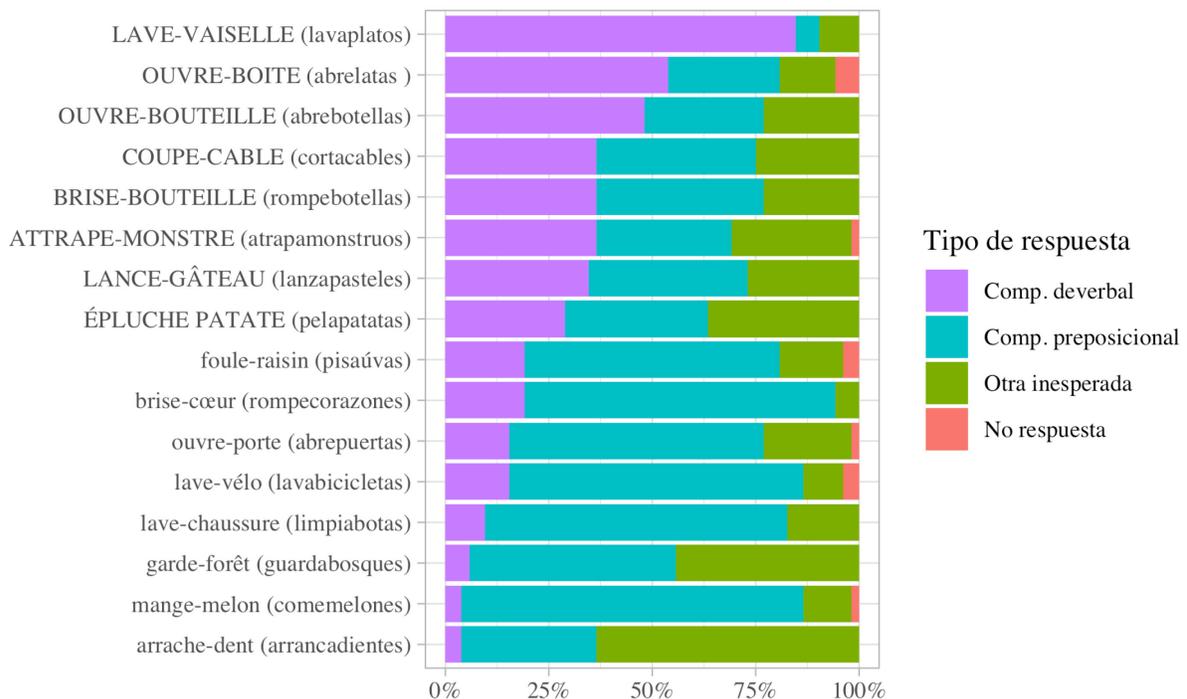


Figura 3.12. Repartición de las respuestas según el ítem en francés.

A través de los ítems, se puede observar una gran diferencia en función del criterio de animacidad (inanimados en mayúscula), con una tasa de 11,54 % de compuestos deverbales en la condición [+animado] y un 44,95 % en la condición [-animado]. Un ítem sobresale en cuanto a la proporción de compuestos deverbales: “*lave-vaisselle*”, probablemente porque se trata de un

compuesto ya muy asentado en francés. El ítem “*arrache-dent*” sobresale también, pero por la proporción de respuestas alternativas que ha elicitado. Como en euskera, una gran mayoría de los participantes contestó con un sustantivo solo (en este caso, “*dentiste*”) en lugar de formar un compuesto deverbal.

Para indagar más en los posibles efectos de la animacidad y de los grupos sobre el tipo de respuesta (1=Compuesto deverbal, 0 = Otros), construimos una regresión logística de efectos mixtos siguiendo los pasos habituales. Es decir, incluimos las variables animacidad y grupo como efectos fijos y también la posible interacción entre ellos (animacidad × grupo). En cuanto a la estructura de los efectos aleatorios, incluimos un intercepto aleatorio por participante e ítem y una pendiente aleatoria por participante para la animacidad. Los factores fueron codificados de la siguiente manera (el nivel de referencia corresponde al primer nivel de codificación): animacidad (-.5=inanimado, .5=animado), grupo (1=EU, 2=FR, 3=FR/EU). Dado que ni el factor grupo ni su interacción con el factor animado mejoraban la bondad de ajuste del modelo, nuestro modelo final solo incluye el factor fijo animado, que es muy significativo ($\beta = -2.869$, $SE = 0.629$, $z = -4.564$, $p < .001$). Esto significa que, en la condición [+animado], todos los grupos producen significativamente menos compuestos deverbales que en la condición [-animado].

En cuanto a la direccionalidad, dentro de las respuestas que corresponden a un compuesto preposicional o propio (el 75,36 % de los datos) y que, por lo tanto, nos permiten evaluar la direccionalidad, el 98,56 % de esas respuestas presentan el elemento verbal al inicio, lo que corresponde al orden canónico del francés. Con una tasa tan alta, no resulta pertinente profundizar más con un análisis estadístico, podemos concluir que los participantes no tienen problemas con la direccionalidad de los compuestos deverbales en francés.

En cuanto a la morfología de estos compuestos, podemos observar en la Figura 3.13 que, cuando un elemento verbal aparece en un compuesto propio, se usa principalmente la conjugación de la tercera persona singular del presente del indicativo (3.PRS.SG), mientras que, cuando aparece en un compuesto preposicional, se usa principalmente el sufijo *-eur/-euse* y, en menor medida, el sufijo *-ien*. Aparecen usos marginales de verbos en infinitivo y verbos flexionados en 3.PRS.SG también en esta condición. Los sufijos *-eur* y *-ien* son ambos válidos para crear un sustantivo a partir del verbo, pero no se pueden usar dentro de un compuesto sin preposición (*“*laveur vaisselle*”, **“gardien forêt”*), sino solo dentro de un compuesto preposicional (“*laveur de*

vaisselle”, “*gardien de forêt*”). El uso mayoritario de *-eur* en lugar de *-ien* se debe principalmente al hecho de que el primero forma derivados a partir de verbos, mientras que el segundo forma derivados a partir de sustantivos (Grevisse & Goosse, 2008, pp. 172-173). Por otro lado, el sufijo *-eur* es más frecuente para formar derivados agentivos que instrumentales, lo que es compatible con el efecto de animacidad que hemos comentado sobre los compuestos deverbales. A su vez, el uso del sufijo *-ien*, que aparece solamente en “*gardien*” proporciona un argumento adicional a favor de lo que hemos comentado anteriormente sobre el carácter nominal de “*garde*” en estas estructuras. En cuanto al uso de la forma infinitiva del verbo en algunos compuestos y sintagmas, podemos estar ante un posible efecto de la prueba, idea que desarrollaremos más en nuestro análisis de las respuestas en español L3.

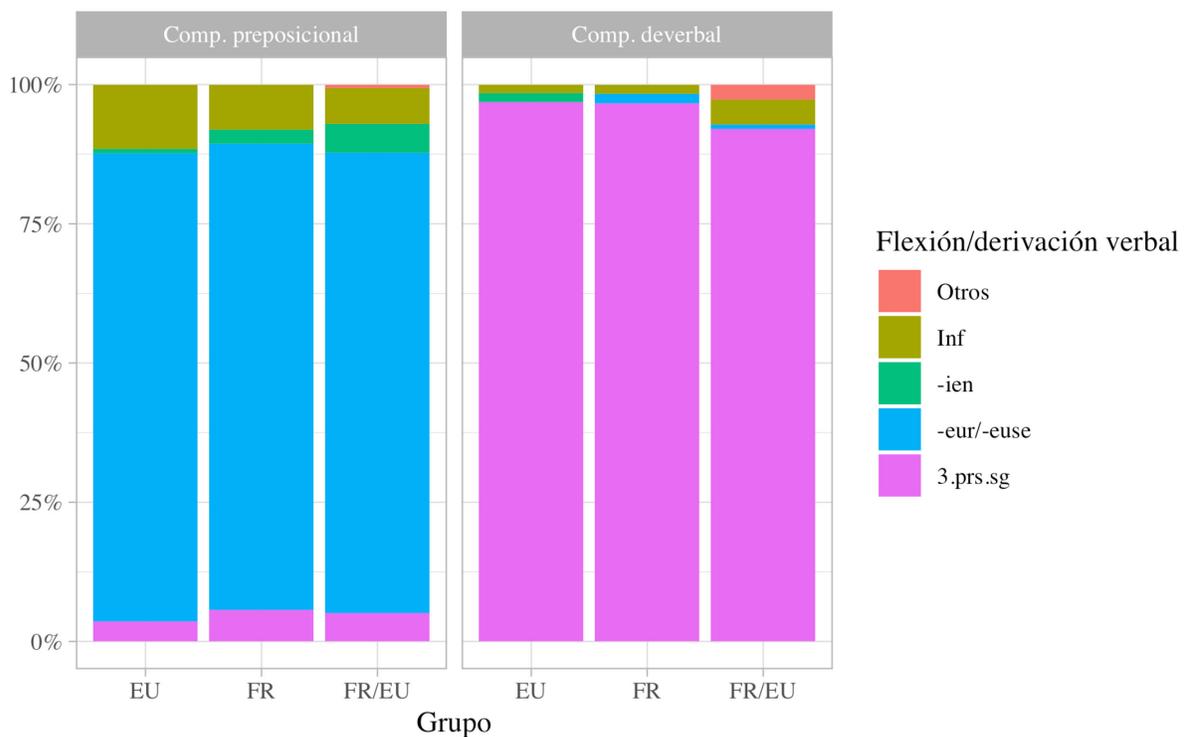


Figura 3.13. Repartición de las respuestas según la morfología de los deverbales, por grupo y estructura.

Con el fin de profundizar el análisis sobre la morfología de los elementos verbales, consideramos solamente las respuestas que contenían verbo flexionado o derivado ($n=688$) y elaboramos un modelo de regresión logística para medir los efectos de grupo y animacidad sobre la respuesta (1=sufijo *-eur/euse*, 0=otros). Empezando con los mismos efectos fijos y aleatorios que en nuestros modelos anteriores y eliminando los efectos que no mejoraban la bondad de ajuste,

llegamos a un modelo que solo incluía el efecto fijo de animacidad y, en cuanto a efecto aleatorios, un intercepto aleatorio por participante e ítem y una pendiente aleatoria por participante para la animacidad. El modelo resultante nos revela que se usa significativamente más el sufijo *-eur/-euse* en la condición [+animado] ($\beta = 2.165$, $SE = 0.893$, $z = 2.425$, $p = .015$). Esto, junto con los resultados de nuestros modelos anteriores nos permite tener una visión más global de la producción de nuestros participantes. En francés, nuestros participantes prefieren el sufijo derivativo *-eur/-euse* para sustantivar el verbo cuando el referente es animado, y la composición deverbal cuando el referente es inanimado. Si bien hemos visto que la tasa de compuestos deverbales solo asciende al 44,95 %, con los referentes inanimados, esto se debe principalmente a otras respuestas alternativas, como un sustantivo solo como los ejemplos de (3.159).

En resumen, todos los grupos de participantes produjeron una mayor cantidad de compuestos preposicionales en francés L1, por lo que los compuestos deverbales propios no parecen ser tan productivos como se indica en la sección 3.2.2.2. A pesar de ello, hemos observado que los compuestos deverbales producidos seguían las normas del francés, es decir, con el elemento verbal en 3SG.PRS al inicio, seguido por su objeto. También hemos observado que el rasgo [-animado] parece favorecer la producción de compuestos deverbales, mientras que el rasgo [+animado] favorece la derivación mediante el sufijo *-eur/-euse*. Este mayor uso de compuestos deverbales para referirse a instrumentos que a agentes coincide con el trabajo de Rosenberg (2008) sobre los compuestos deverbales del francés.

3.7.2.3 Español L1

Antes de explorar los datos en la tercera lengua de nuestros grupos experimentales, analizamos la producción de nuestro grupo de control en su primera lengua, el español.

Como en el caso de los compuestos N+N, este análisis nos permite comprobar la validez de la prueba en español, ya que las respuestas elicitadas ($n=336$) corresponden principalmente a compuestos deverbales (87,50 %) que presentan el orden V+N (100 %), la flexión verbal de 3.SG.PRS (99,31 %) y la marca de la -s genérica (98,68 %). No se observa ningún patrón específico en cuanto a los pocos errores de flexión o de marca de la -s, por lo que atribuimos estas idiosincrasias a errores de inatención por parte de los participantes. Sin embargo, dada la variación observada en el tipo de respuestas producidas, profundizamos nuestro análisis al respecto.

Aparte de los compuestos deverbales propios, que corresponden a 87,50 % de la producción total, los participantes también produjeron compuestos preposicionales con un derivado verbal (3,57 %), como en (3.162). También encontramos un 8,63 % de otras formas inesperadas, pero ningún caso de no respuesta. Dentro de la categoría “compuestos deverbales”, hemos incluido también respuestas que, además del compuesto deverbal propio, contenían otros elementos (por ejemplo, lo que sería el sujeto del verbo) cuando no afectaban al significado del compuesto, como muestran los ejemplos (3.163). En estos casos, el compuesto deverbal es recategorizado como un adjetivo, pero no esto no altera su estructura interna.

- (3.162) a. un pisador de uvas / un pisaúvas
 b. un abridor de latas / un abrelatas
 c. un lavador de bicis / un lavabicycletas
- (3.163) a. un señor rompecorazones / un rompecorazones
 b. un señor abrepuertas / un abrepuertas

En cuanto a las formas alternativas que no corresponden a un compuesto del tipo Objeto-Verbo o Verbo-Objeto, encontramos una variedad de respuestas. Como en euskera y en francés, la mayoría de las respuestas alternativas consisten en un sustantivo solo (7,48 %) que conlleva una gran parte del significado del compuesto que queríamos elicitare, como muestran los ejemplos de (3.164). Otras respuestas corresponden simplemente a un derivado verbal solo (10 %), como en (3.165). Encontramos también un caso de estructura N+para+N, que reproducimos en el ejemplo (3.166).

- (3.164) a. martillo / rompebotellas
 b. dentista / arrancadientes
 c. trampa / atrapamonstruos
 d. catapulta / lanzapasteles
 e. portero / abrepuertas
- (3.165) a. pelador / pelapatatas
 b. lavadero / lavabicycletas
- (3.166) trampa para osos / atrapamonstruos

Tras esa breve exploración de las respuestas inesperadas, volvemos a la distribución global de los compuestos deverbales propios (87,50 %) y las formas alternativas. Como en nuestros análisis anteriores, nos interesa el impacto de los rasgos [+animado] y [+atestiguado] sobre la producción de respuestas esperadas. En cuanto a la animacidad, la tasa de compuestos deverbales es del 86,49 % en la condición [+animado] y del 88,09 % en la condición [-animado]. En cuanto a

la diferencia entre los ítems atestiguados o no, la tasa de compuestos deverbales asciende al 96,42 % en la condición [+atestiguado] y baja al 78,57 % en la condición [-atestiguado].

Para evaluar la significatividad de estos efectos sobre el tipo de respuesta (1=V+N; 0=Otros), aplicamos un modelo de regresión logística de efectos mixtos que incluía las variables animado y atestiguado como efectos fijos y también la posible interacción entre ellos (animado \times atestiguado). En cuanto a la estructura de los efectos aleatorios, incluimos un intercepto aleatorio por participante e ítem y una pendiente aleatoria por participante para las variables animado y atestiguado. Los factores fueron codificados de la siguiente manera (el nivel de referencia corresponde al primer nivel de codificación): animacidad (-.5=inanimado, .5=animado), atestiguado (-.5=no atestiguado, .5=atestiguado). Dado que ni el factor animado ni su interacción con el factor atestiguado mejoraban la bondad de ajuste del modelo, nuestro modelo final solo incluye el factor fijo atestiguado ($\beta = 2.114$, $SE = 1.048$, $z = 2.016$, $p = .044$). El valor positivo de este coeficiente y su significatividad nos confirman que los ítems que representan un compuesto atestiguado favorecen la producción de compuestos deverbales, lo que es lógico dado que estos compuestos se encuentran más a menudo en el *input* de los hablantes nativos.

Más allá de la diferencia a nivel global, a través de los ítems, se puede observar cierta variación en cuanto al hecho de que un compuesto esté atestiguado o no. La Figura 3.14 ilustra las diferentes formas producidas para cada ítem, los ítems atestiguados aparecen en mayúsculas.

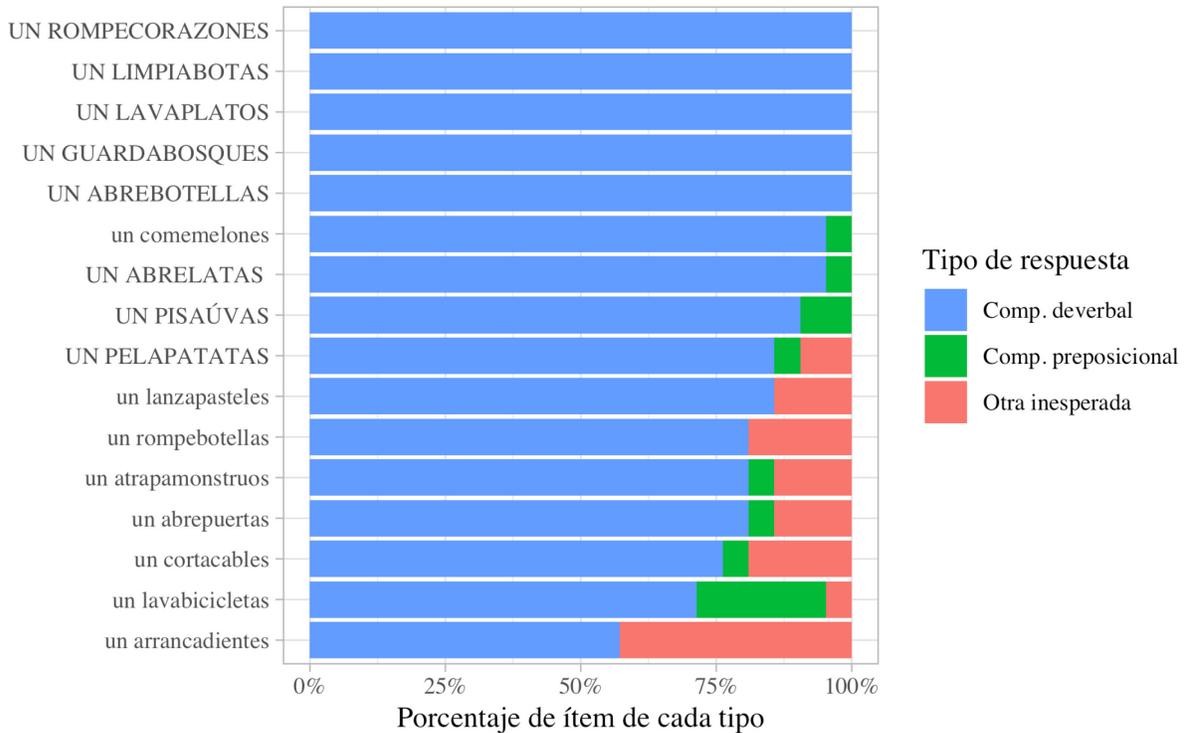


Figura 3.14. Repartición de los tipos de respuestas según los ítems de compuestos deverbales en español L1.

De los ítems atestiguados, “pelapatatas” sobresale con una tasa un poco más alta de formas alternativas. Estas son “pelador” y “pelador de patatas”, dos opciones válidas para el significado esperado. Esta pista nos lleva a pensar que el factor más relevante para que se produzca o no un compuesto deverbale podría estar vinculado con otros dos factores: (1) la existencia de una forma equivalente más sencilla, o (2) el uso frecuente o atestiguado del compuesto deverbale para otro significado que el que queríamos elicitar en nuestra prueba. El caso de “pelador”, que coexiste con el vocablo bien asentado “pelapatatas” quizá no sea el más llamativo, pero si tomamos, por ejemplo, los casos de “arrancadientes”, “atrapamonstruos” o “abrepuertas” y sus alternativas “dentista”, “trampa” y “portero”, aparece de manera clara que no es necesario recurrir a un compuesto deverbale, ya que existe una opción aproximada más breve y sencilla. En cambio, con el ítem para elicitar “lavabicycletas”, el uso de un sintagma preposicional (“lavador de bicycletas”, “limpiador de bicycletas”) podría estar relacionado con el hecho de que este compuesto, si bien no está atestiguado, parece estar relacionado con un uso más instrumental/inanimado (quizá por analogía con los compuestos deverbales “lavavajillas”, “lavaplatos” que hacen referencia a una máquina y no a una persona). Con el objetivo de indagar

más en este orden de ideas, recurrimos al *Corpus del Español: Now* (CdEN; Davies, 2019) para averiguar la frecuencia de los sustantivos y de los compuestos sintagmáticos preposicionales que aparecían en las respuestas de nuestros participantes.¹⁸⁰ La Tabla 3.11 presenta las respuestas alternativas de nuestros participantes, junto con su frecuencia absoluta en el CdEN. Los compuestos deverbales que hemos considerado como atestiguados en el diseño de nuestro estudio aparecen en mayúsculas.

Tabla 3.11. Respuestas esperadas y alternativas de los participantes, acompañadas de sus frecuencias en el CdEN.

Comp. deverbales esperado	Frec.	Sustantivo alternativo	Frec.	Comp. preposicional alternativo	Frec.
abrepuestas	96	portero	221412	abridor de puertas	6
atrapamonstruos	0	trampa	54950	trampa para N ⁱ	449
rompebotellas	0	martillo	21221		
arrancadientes	0	dentista	12082		
lanzapasteles	0	catapulta	2745		
cortacables	20	alicates/cortador	2381	cortador de N	799
lavabicycletas	0	lavador	1216	limpiador de N lavador de N	1437
PELAPATATAS/PELAPAPAS	56	pelador	322	pelador de N	96
PISAÚVAS	1			pisador de N	8
comemelones	0			comedor de N ⁱⁱ	191
ABRELATAS	509			abridor de lata(s)	19
GUARDABOSQUES	6684				
LIMPIABOTAS	782				
ROMPECORAZONES	994				
ABREBOTELLAS	64				
LAVAPLATOS	1754				

ⁱ Como para varios compuestos preposicionales de nuestro juego de datos, no hemos encontrado ocurrencias precisas de “trampa para monstruos”, pero hemos encontrado varias ocurrencias de “trampa+para+N”. Reportamos la frecuencia de esta estructura más general, donde N equivale a cualquier sustantivo.

ⁱⁱ Comedor de melones no tenía ninguna ocurrencia en el corpus, comedor de sandías tenía dos. Al ampliar la búsqueda a “comedor+de+N”, constatamos que la mayoría de las entradas hacían referencia a un lugar, por lo que hemos filtrado los resultados para conservar solamente las entradas en las que “comedor” tenía un valor agentivo (p. ej. “comedor de pescado”, “comedor de plantas”, etc.).

¹⁸⁰ A la hora de hacer la búsqueda, el CdEN contaba con aproximadamente 7200 millones de palabras extraídas a partir de diferentes páginas web de periódicos y revistas de todo el mundo hispano, desde el año 2012 hasta la actualidad. El corpus sigue creciendo cada mes.

Esta tabla nos revela que las producciones alternativas de nuestros participantes no son tan idiosincrásicas. Efectivamente, los nueve sustantivos y los diez compuestos preposicionales alternativos utilizados por nuestros participantes aparecen, por lo menos una vez, en el CdeN. La frecuencia de cada expresión en el corpus varía de manera importante, por lo que intentamos comparar la frecuencia dentro del corpus con la frecuencia de las respuestas alternativas en nuestros datos.

La Figura 3.15 ilustra las clasificaciones del tipo de respuesta por ítem, en orden descendiente según la frecuencia de la forma alternativa en el corpus. Con una inspección visual de la figura, no vemos una correlación clara entre la frecuencia en el corpus y la frecuencia en la producción de las formas alternativas por parte de nuestros participantes, pero es posible que esto se deba al tamaño reducido de nuestro juego de datos. Una vez separadas por tipo de respuesta (sustantivo solo, compuesto preposicional), la baja cantidad de formas alternativas producidas por nuestros participantes no nos permite conducir ninguna comparación estadística relevante.

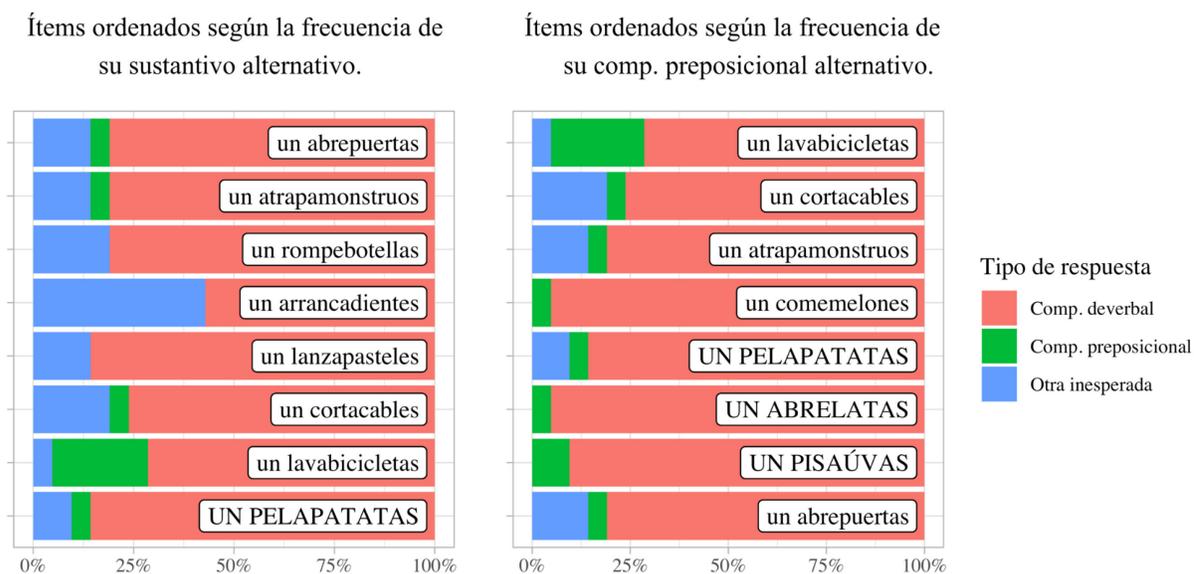


Figura 3.15. Repartición del tipo de respuesta según los ítems deverbales en español L1, ordenados por frecuencia de las alternativas en el CdEN (solo aparecen los ítems que tenían formas alternativas).

Sin embargo, a partir de los datos de nuestra prueba y de los datos del corpus, podemos confirmar, por una parte, que, cuando existen alternativas a los compuestos deverbales, ciertos hablantes prefieren esas alternativas. Por otra parte, nuestros datos nos confirman que los

compuestos deverbales son muy productivos en español (como sostiene la literatura, e.g. Tuggy 2003 y Val Álvaro 2000) y representan la opción más frecuente para acuñar nuevas expresiones cuando no existe ningún vocablo equivalente.

3.7.2.4 Español L3: las etapas iniciales

Antes de continuar con los datos del español como L3, resulta necesario ajustar nuestras hipótesis conforme a lo que hemos descubierto sobre los compuestos deverbales en las dos primeras lenguas de nuestros participantes. En cuanto a la productividad de los compuestos deverbales en esas dos lenguas, hemos visto que la producción de compuestos deverbales en francés era mucho menor que la de sus equivalentes en euskera.

A modo de recapitulación, presentamos en la Tabla 3.12 las tasas de productividad de los compuestos deverbales y su direccionalidad en las lenguas nativas o casi nativas de los participantes. Para cada lengua, dividimos la tasa de productividad en función del rasgo [\pm animado], que tuvo un impacto significativo sobre la productividad de los compuestos deverbales en francés.

Tabla 3.12. Productividad y direccionalidad de los compuestos deverbales en las L1/L2 de los participantes.

	Euskera		Francés		Español	
	+	-	+	-	+	-
Animado	+	-	+	-	+	-
Producción de compuestos deverbales	81,25 %	72,36 %	11,54 %	44,95 %	86,49 %	88,09 %
Direccionalidad canónica	99,22 %		98,56 %		100 %	

A la luz de estos datos, podríamos reformular la hipótesis D1 en dos versiones, según la fuente de influencia:

- D1-FR: Los participantes que muestren una influencia del francés deberían producir un gran número de compuestos sintagmáticos preposicionales como “limpiador de botas”, en lugar de compuestos deverbales propios como “limpiabotas”.
- D1-EU: Los participantes que muestren una influencia del euskera deberían producir una mayoría de compuestos deverbales sin nexo, como “limpiabotas”, “limpiador botas”, “botas limpiador” o

“botas limpia” y evitar los compuestos sintagmáticos preposicionales como “limpiador de botas”.

Las demás hipótesis se mantienen inalteradas, pero las repetimos aquí a modo recordatorio:

- D2-FR: Los participantes que muestren una influencia del francés deberían producir compuestos deverbales con la direccionalidad adecuada, es decir, el deverbale primero, seguido por un sustantivo.
- D2-EU: Los participantes que muestren una influencia del euskera deberían producir compuestos deverbales con la direccionalidad inversa, es decir, un sustantivo seguido por un deverbale.
- D3-FR: Los participantes que muestren cierta influencia del francés deberían producir menos compuestos deverbales con los ítems animados.
- D3-EU: El rasgo de animación no debería influir sobre la productividad de los deverbales para los participantes que muestren cierta influencia del euskera. Si se observa un cambio en función de este rasgo, el cambio debería afectar la forma y no la productividad como tal.
- D4: Al tratarse de unidades léxicas presentes en la lengua y, por lo tanto, susceptibles de formar parte del *input* recibido por los aprendientes, los ítems atestiguados deberían elicitar una mayor tasa de respuestas esperadas (V+N) para todos los participantes.
- D5-FR: Los participantes que muestren una influencia del francés deberían producir compuestos deverbales con un deverbale flexionado correctamente (en tercera persona del singular).
- D5-FR-alt: En las primeras recolecciones de datos, los participantes que muestren una influencia del francés deberían usar formas idiosincrásicas para el elemento deverbale, es decir, o el verbo en infinitivo como “limpiarbotas”, o una flexión verbal alternativa como “limpiobotas” (sobre todo la primera persona del singular según Salomaa-Robertson, 2000). El uso de esas formas idiosincrásicas debería decaer con el tiempo y una mayor exposición a la lengua, cuando los aprendientes vayan dominando la flexión verbal en presente del indicativo.
- D5-EU: Los participantes que muestren una influencia del euskera deberían derivar el deverbale a partir de un proceso similar a los deverbales del euskera, es decir, con un sufijo agentivo o instrumental que podría corresponder a -dor(a) en español.
- D6: Los participantes que muestren una influencia de cualquiera de sus dos primeras lenguas deberían, por lo menos al principio, producir los compuestos deverbales sin la -s final. Con una mayor experiencia con la lengua y un mayor *input*, la -s final debería aparecer en los datos de las segunda y tercera recolección.

Con estas hipótesis en mente, continuamos el análisis con nuestros grupos experimentales en español L3. La primera sección de presentación de resultados cubre las hipótesis D1 a D4 y la segunda mitad corresponde a las hipótesis D5 y D6.

3.7.2.4.1 *Productividad y direccionalidad*

En cuanto al tipo de respuestas proporcionadas, los datos son más heterogéneos que en las lenguas nativas o casi nativas, y, si bien corresponden principalmente a compuestos deverbales (75,00 %), encontramos una tasa más elevada que en los casos anteriores tanto de respuestas alternativas (18,50 %) como de no respuesta (6,49 %). Dentro de estas respuestas alternativas, al contrario de los datos en francés, los compuestos preposicionales que incluyen un verbal son muy escasas (2,04 %). En cuanto a las formas alternativas que no corresponden a un compuesto del tipo Objeto-Verbo o Verbo-Objeto, encontramos una variedad de respuestas. Como en los resultados anteriores de las tres lenguas, algunas respuestas consisten en un sustantivo solo que integra buena parte del significado del compuesto que queríamos elicitare, como se ve en los ejemplos de (3.167). Otras respuestas corresponden simplemente a un derivado verbal solo, como muestra el ejemplo (3.168) o la forma N+N/N+de+N, como vemos en los ejemplos de (3.169).

- | | | | |
|------------|----------------|------------------|-----------------|
| (3.167) a. | policía | / | guardabosques |
| | b. | catapulta | / lanzapasteles |
| (3.168) | lavadores | / | lavaplatos |
| (3.169) a. | policía bosque | / | guardabosques |
| | b. | hombre de bosque | / guardabosques |

Los resultados globales, en comparación con el grupo de español L1, aparecen en la Figura 3.16 (gráfico de la izquierda). Con respecto a la direccionalidad (Figura 3.16, gráfico de la derecha), se observa un efecto de techo para todos los grupos (100 % para el grupo de control, 98,1 % para los grupos experimentales), por lo que no es pertinente la elaboración de modelos de análisis estadístico para este factor.

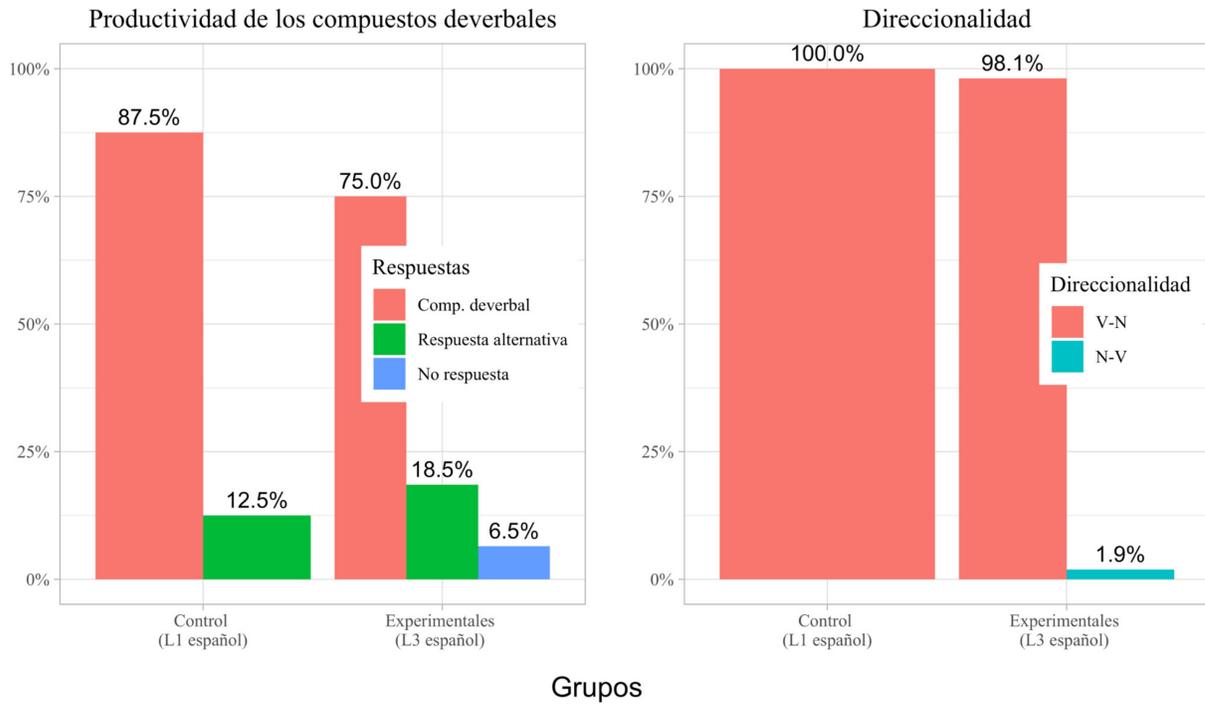


Figura 3.16. Distribución de las respuestas para los ítems de compuestos deverbales en español. Todos los grupos experimentales (EU, FR, FR/EU) aparecen juntos.

Siguiendo los mismos pasos que para los dos modelos anteriores, realizamos una regresión logística de efectos mixtos para la productividad (1=VN; 0=Otros) de los compuestos deverbales. Los efectos fijos del modelo resultante para la productividad de los VN aparecen en la Tabla 3.13. Además de los interceptos aleatorios por participante e ítem, incluimos también una pendiente aleatoria por participante para las variables animado y atestiguado. En cuanto al factor grupo, dado que comparamos los grupos experimentales con el grupo de control, hemos elegido la codificación siguiente, con el primer nivel como referencia: grupo (1=ES, 2=EU, 3=FR, 4=FR/EU).

Tabla 3.13. Resultados del modelo de regresión logística de efectos mixtos del tipo de respuesta esperada para los ítems de compuestos deverbales en español.

	β	<i>SE</i>	<i>z</i>	<i>Pr(> z)</i>
(Intercepto)	4.408	0.767	5.749	<.001
Animado	-1.851	0.609	-3.039	.002
Atestabilidad	2.963	0.800	3.703	<.001
Grupo (EU)	-2.128	0.942	-2.260	.024
Grupo (FR)	-2.097	0.937	-2.238	.025
Grupo (FR/EU)	-1.921	0.859	-2.237	.025
Atestiguado × Grupo (EU)	-2.959	0.830	-3.564	<.001
Atestiguado × Grupo (FR)	-3.890	0.873	-4.455	<.001
Atestiguado × Grupo (FR/EU)	-3.309	0.737	-4.491	<.001

Si bien todos los efectos son significativos y todos los grupos experimentales producen significativamente más respuestas alternativas que el grupo de control, un análisis posterior del mismo modelo, sin los datos del grupo de control, nos revela que solo el efecto de animacidad es significativo para los grupos experimentales ($\beta = -2.589$, $SE = 1.003$, $z = -2.581$, $p = .009$). Dicho de otro modo, el factor “atestiguado” solo afecta de manera significativa al grupo de control, mientras que los grupos experimentales solo se ven afectados por el efecto de animacidad. Una prueba *posthoc* de comparaciones múltiples de medias (Contraste Tukey) no nos revela ninguna diferencia significativa entre los grupos experimentales (FR – EU: $\beta = 0.032$, $SE = 0.977$, $z = 0.032$, $p = 0.97$; FR/EU – EU: $\beta = 0.206$, $SE = 0.906$, $z = 0.228$, $p = 0.82$; FR/EU – FR: $\beta = 0.175$, $SE = 0.871$, $z = 0.201$, $p = .84$). Con este modelo podemos confirmar que los participantes de todos los grupos experimentales son más propensos a producir compuestos deverbales cuando se trata de un referente inanimado (80,52 %) que cuando es animado (69,47 %).

3.7.2.4.2 Morfología de los compuestos deverbales

Como vimos en la sección 3.2.1.2, los compuestos deverbales muestran una serie de particularidades en lo que respecta a la morfología, tanto del verbo como del sustantivo que le acompaña. En el caso del verbo, lo esperado es que aparezca flexionado en tercera persona del singular, pero en las respuestas de nuestros grupos experimentales aparece una variedad de formas, ilustradas en (3.170). Dada la escasez de errores del grupo de control, esta sección se concentra únicamente en los resultados de los grupos experimentales.

(3.170) a. abrebotellas [3SG.PRS, respuesta esperada]

- b. *abrobotellas [1SG.PRS]
- c. *abrirbotellas [infinitivo]
- d. *abribotellas [otro]

Como podemos ver en la Figura 3.17, la morfología esperada (3SG.PRS) aparece entre las formas producidas, pero no de forma mayoritaria; la forma más frecuente en los datos es el verbo en infinitivo. Con respecto al comportamiento de los distintos grupos de participantes, el grupo FR es el que produce la flexión verbal adecuada en menor proporción.

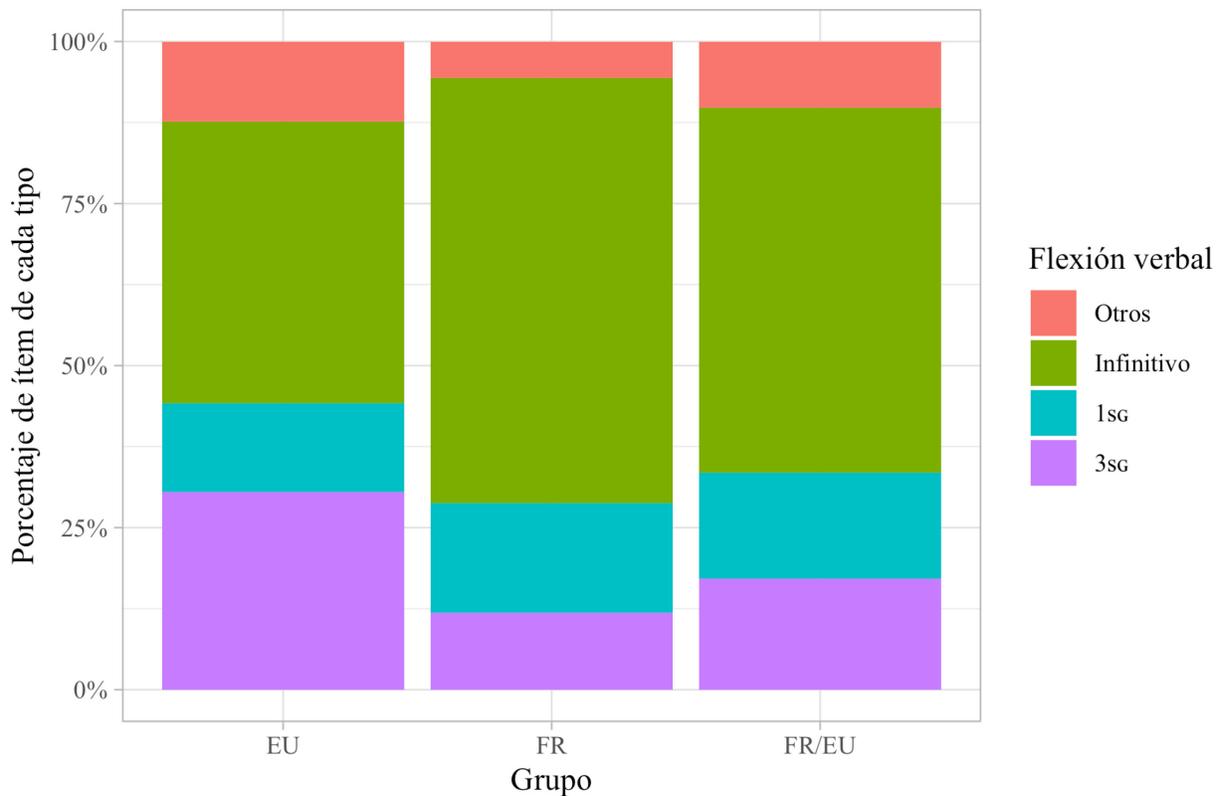


Figura 3.17. Morfología del verbo de los compuestos deverbales (n=624) - Grupos experimentales.

En ninguno de los compuestos VN se deverbaliza el verbo mediante un sufijo -dor u otros sufijos agentivos o instrumentales parecidos, lo que parece indicar una ausencia de influencia del euskera, que, como el inglés, requiere un sufijo derivativo para crear un deverbal.¹⁸¹ Si bien los

¹⁸¹ En los resultados aparece un único uso de sufijo derivativo, que aparece formando parte del compuesto sintagmático preposicional que reproducimos en (i), que representa una estructura aceptable tanto en francés como en español. Sin

datos no muestran una influencia del euskera en cuanto a la derivación del verbo, tampoco queda claro cuál es el peso de la influencia del francés. Como hemos visto, los participantes producen una mayoría de formas idiosincrásicas, como un verbo en infinitivo o un verbo flexionado en la primera persona del singular, en lugar de emplear la tercera persona del singular esperada en los deverbales tanto del español como del francés. Una posible explicación de estos resultados puede ser una simple falta de dominio de la flexión verbal del español, al tratarse de participantes que se encuentran en las etapas iniciales de adquisición del español. Por otro lado, un análisis detallado de los ítems también arroja una serie de pistas para comprender la variedad de formas verbales producidas.

En primer lugar, encontramos lo que parece ser un efecto de la tarea. Ciertos ítems presentaban en el estímulo el verbo conjugado en primera persona, como vemos en (3.171), mientras que otros usaban una perífrasis con los verbos “querer” o “gustar” que mostraba el verbo meta en infinitivo, como muestran los ejemplos (3.172) y (3.173).

- (3.171) Como dos melones cada día. Soy un... (respuesta esperada: comemelones)
- (3.172) Quiero pelar una patata. Necesito un... (respuesta esperada: pelapatatas)
- (3.173) Me gusta mucho limpiar las botas. Soy un... (respuesta esperada: limpiabotas)

Los participantes lograron producir compuestos con la flexión de tercera persona del singular en ambos tipos de ítem, aunque en proporciones variables. Además, si agrupamos los resultados por tipo de estímulo (Figura 3.18), resulta evidente que los estímulos con un verbo conjugado en primera persona favorecen los deverbales con esa misma forma verbal, mientras que los ítems con perífrasis favorecen el uso del infinitivo.

embargo, en español, esa respuesta tiene una forma idiosincrásica que no procede de la estructura del sintagma sino de la derivación errónea (-or en lugar de -dor). El equivalente en francés aparece debajo del ejemplo.

(i) un *limpior de botas
un laveur de bottes

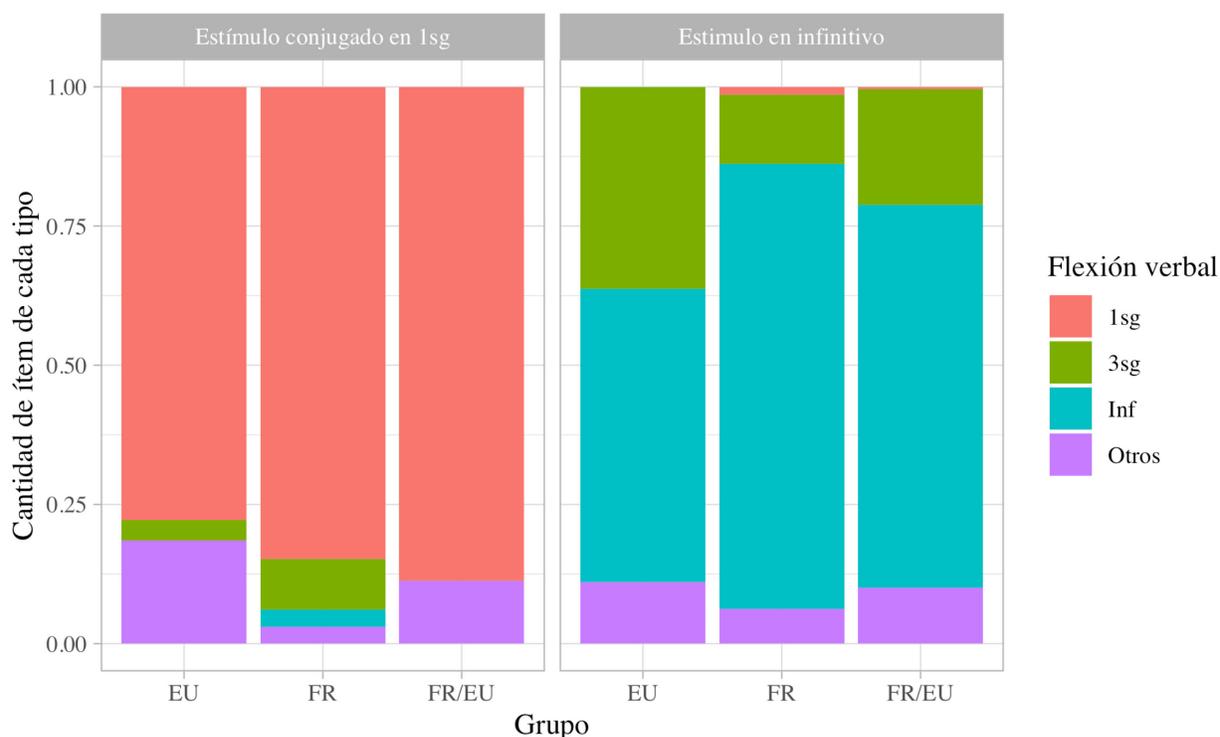


Figura 3.18. Flexión verbal de los compuestos deverbales en español L3, según la paráfrasis del estímulo.

Siguiendo el mismo procedimiento que en la sección anterior, aplicamos un modelo de regresión logística de efectos mixtos para evaluar la flexión verbal (3SG.PRS = 1, Otros = 0) con los efectos fijos de estímulo (Conjugado = 0, Infinitivo = 1), grupo (EU = 1, FR = 2, FR/EU = 3) y la interacción estímulo \times grupo. En cuanto a los efectos aleatorios, solo logramos incluir un intercepto aleatorio por participante e ítem. Este análisis nos revela que el efecto de estímulo es significativo ($\beta = 5.977$, $SE = 2.432$, $z = 2.458$, $p = .014$) y que interactúa de manera significativa con el efecto de grupo, pero solo para el grupo FR ($\beta = -5.819$, $SE = 1.723$, $z = -3.387$, $p < .001$). El efecto principal de grupo no es significativo (FR: $\beta = 2.551$, $SE = 2.313$, $z = 1.103$, $p = 0.26$, FR/EU: $\beta = -18.654$, $SE = 178.259$, $z = -0.105$, $p = .92$).¹⁸²

Si bien los resultados apuntan a que los estímulos con el verbo meta en infinitivo favorecen el uso de un verbo flexionado en algunos casos, cabe señalar que la forma favorecida no parece ser exactamente la flexión de tercera persona del singular, sino más bien una apócope de la -r del verbo

¹⁸² El modelo calcula coeficientes exagerados para el grupo FR/EU porque ese grupo no produce ninguna flexión verbal de 3SG.PRS con los estímulos que contenían un verbo conjugado.

en infinitivo. El indicio de esta posibilidad nos la dan los verbos de tercera conjugación (-IR). Los verbos de la primera y segunda conjugación, como “lavar” y “romper”, elicitan “lava” y “rompe”, respectivamente, que es compatible con las dos formas (3SG.PRS e infinitivo apocopado). Por su parte, los verbos de la tercera conjugación, como “abrir”, elicitan “*abri” en lugar de “abre”. De manera crucial, para los estímulos con un verbo de la tercera conjugación, no encontramos ningún caso de flexión verbal correcta, sino formas acabadas en -i, lo que parece confirmar esa estrategia de apócope del infinitivo. Al mismo tiempo, estos resultados sugieren que nuestros participantes no dominan la flexión verbal en estas primeras etapas de adquisición de su uso del español, efecto que también ha sido observado en el trabajo de Salomaa-Robertson (2000) sobre los compuestos en español L2. Los resultados detallados para cada conjugación aparecen en la Tabla 3.14.

Tabla 3.14. Flexión de los deverbales para todas las respuestas de tipo V+N a un estímulo en infinitivo.

	1ª conjugación (-AR)	2ª conjugación (-ER)	3ª conjugación (-IR)
3sg	93	21	0
Infinitivo	215	54	78
Otros	1	0	46
Total	309/416 ⁱ	75/104	124/156

ⁱ En esta línea, la cifra a la derecha de la barra diagonal representa el total de respuestas V+N cuyo estímulo correspondía a la conjugación correspondiente.

Con respecto a la -s genérica del sustantivo de los deverbales, encontramos nuevamente un posible efecto de la tarea, es decir, los estímulos que presentaban el complemento del verbo en plural, como en los ejemplos (3.171) y (3.173), favorecieron la producción de la -s genérica en las respuestas V+N, mientras que la formulación en singular favoreció la producción de V+N sin -s. La Tabla 3.15 ilustra el contraste entre los dos tipos de estímulos.

Tabla 3.15. Repartición de las respuestas con o sin -s genérica en función de la paráfrasis del estímulo.

	Paráfrasis en singular	Paráfrasis en plural
Ausencia de la -s	305	29
Presencia de la -s	3	304

Si dejamos de lado los resultados de los ítems con el estímulo en plural para intentar controlar el posible efecto del estímulo, la escasa frecuencia de producción esperada (sustantivo en plural) nos muestra que todos los grupos experimentales tienen la misma dificultad con la pluralización del segundo elemento del compuesto de verbal, algo que también observó Salomaa-

Robertson (2000), incluso en los aprendientes de español de nivel intermedio y, en menor medida, con los de nivel avanzado.¹⁸³

A continuación, exploramos los datos longitudinales para ver cómo evoluciona la producción de compuestos deverbales de nuestros participantes a medida que van progresando en español.

3.7.2.5 Español L3: perspectiva longitudinal

Para ver cómo evoluciona el dominio del fenómeno con el paso del tiempo en nuestros grupos experimentales, comparamos los datos de la primera recolección con los de las dos recolecciones sucesivas que se llevaron a cabo después de 140 y 200 horas de enseñanza en español, respectivamente. Nos concentraremos en los aspectos de más interés que hemos observado en las etapas iniciales, es decir, la tasa de producción compuestos deverbales en función de los grupos (D1) y en función de los rasgos animado (D3) y atestiguado (D4), la tasa de aciertos en cuanto a la flexión verbal (D5) y a la -s genérica en el segundo componente del deverbale (D6).

Como hicimos en el análisis longitudinal de los compuestos N+N, para comprobar el efecto de cada variable, construimos un modelo para cada recolección de datos con todos los factores relevantes siguiendo los pasos de los modelos de la primera recolección, con la misma prueba *posthoc* para determinar las diferencias entre los grupos experimentales. Además, construimos un modelo global para las tres recolecciones que incluía el factor recolección para comprobar la evolución en el tiempo.

En cuanto a la productividad de los compuestos deverbales en las respuestas, la tasa de producción de esa estructura en nuestros grupos experimentales pasa de 75,00 % en la primera recolección, a 78,9 % en la segunda (aumento significativo: $\beta = 0.314$, $SE = 0.138$, $z = 2.298$, $p = 0.022$) y a 94,49 % en la tercera recolección (aumento significativo también: $\beta = 1.928$, $SE = 0.195$, $z = 9.905$, $p < .001$). En la segunda recolección, la diferencia entre los grupos experimentales se mantiene, el grupo FR produce significativamente menos compuestos deverbales que el grupo EU ($\beta = -2.8767$, $SE = 1.4312$, $z = -2.010$, $p = .044$), mientras que el grupo FR/EU solo se diferencia

¹⁸³ Recordamos al lector que la vulnerabilidad de la -s final en los compuestos deverbales se observó también con tareas de decisión léxica cronometrada (Desrochers et al., 2003) y de juicios de gramaticalidad (Liceras et al., 2004) con aprendientes de español L2/L3 de todos los niveles y, en menor medida, con hablantes nativos del español.

del grupo EU de manera marginalmente significativa ($\beta = -2.4617$, $SE = 1.2992$, $z = -1.895$, $p = .058$). En la tercera recolección, la diferencia entre los grupos experimentales deja de ser significativa (FR: $\beta = -0.3765$, $SE = 0.8332$, $z = -0.452$, $p = .651$; FR/EU: $\beta = 1.5132$, $SE = 0.9194$, $z = 1.646$, $p = .010$). En esta última recolección de datos, los participantes de los grupos experimentales sobrepasaron al grupo de control en la tasa de producción de compuestos deverbales (87,50 %), lo que demuestra que nuestros participantes son conscientes de la productividad del proceso y lo aplican de manera más sistemática incluso que los participantes del grupo de control, que en ocasiones recurren a formas alternativas (por ejemplo, “pelador” en lugar de “pelapatatas”). En cuanto al efecto de la animacidad sobre el tipo de compuesto producido, la significatividad de dicho efecto solo es marginal a partir de la segunda recogida ($\beta = -1.085$, $SE = 0.640$, $z = -1.694$, $p = .093$), posiblemente a medida que los aprendientes se van dando cuenta de que el proceso de composición en español se puede usar para crear compuestos tanto agentivos como instrumentales. El efecto del rasgo atestiguado no mejora la bondad de ajuste en ninguno de los modelos, por lo que concluimos que no es significativo en ninguna de las tres recogidas para los grupos experimentales.

En cuanto a las características morfológicas de los deverbales, los participantes de los grupos experimentales se siguen encontrando lejos de los nativos incluso en la última recolección. El gráfico de la izquierda en la Figura 3.19 muestra que la tasa de aciertos con la marca del plural en el sustantivo no evoluciona de manera significativa con el tiempo. El efecto de la tarea observado en la primera recolección se mantiene a lo largo de las tres recolecciones, es decir, cuando el objeto aparece en plural en el estímulo, se favorece la marca del plural en el compuesto verbal. No es sorprendente, entonces, que el promedio de presencia de esta marca se mantenga en torno al 50 %, ya que la mitad de los ítems contiene una paráfrasis en plural y la otra mitad en singular.

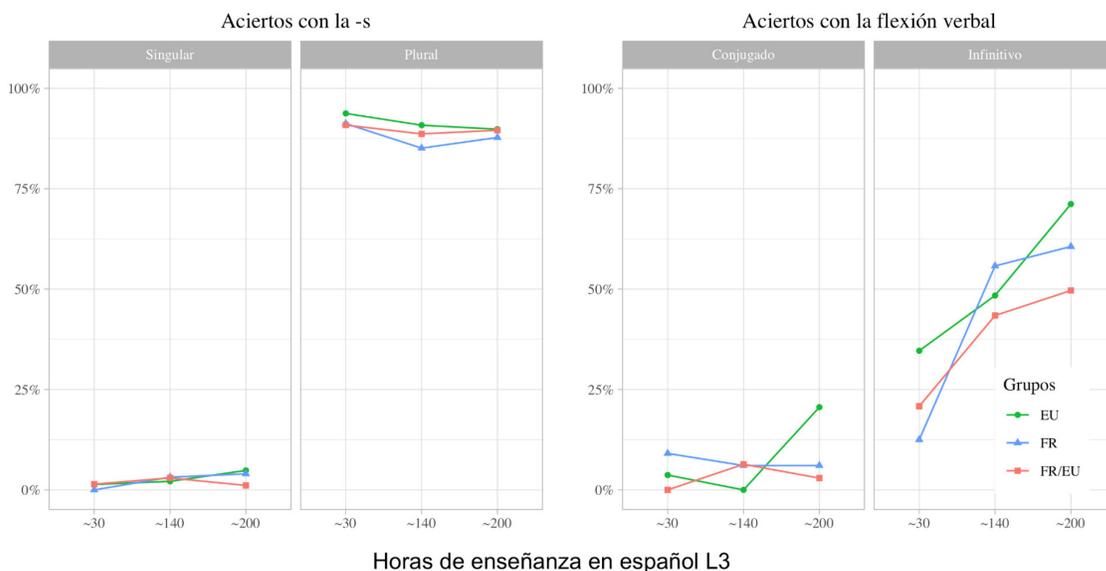


Figura 3.19. Evolución de los aciertos en cuanto a la -s genérica y la flexión verbal en los compuestos deverbales.

En cambio, como vemos en el gráfico de la derecha de la Figura 3.19, la tasa de aciertos con la flexión verbal sí muestra una evolución significativa entre la primera y la segunda recogida ($\beta = 2.129$, $SE = 0.198$, $z = 10.732$, $p < .001$) y también entre la segunda y la tercera etapa de la recogida ($\beta = 1.182$, $SE = 0.181$, $z = 6.182$, $p < .001$), todo ello a pesar del efecto del estímulo (Conjugado vs. Infinitivo), que también tiene un efecto significativo global sobre las tres recolecciones de datos ($\beta = 3.734$, $SE = 1.701$, $z = 2.195$, $p = .028$). El efecto de grupo solo es significativo en la tercera etapa entre los grupos FR/EU – EU ($\beta = -3.884$, $SE = 1.414$, $z = -2.747$, $p = .006$) y FR – EU ($\beta = -3.5503$, $SE = 1.5619$, $z = -2.273$, $p = 0.023$). A pesar de la evolución, los grupos experimentales terminan con un promedio de aciertos del 49,68 % en cuanto a la flexión verbal y de apenas el 21,53 % si consideramos la flexión verbal y la marca del plural de manera conjunta. Podemos concluir, entonces, que los participantes están todavía lejos de dominar los procesos morfológicos que subyacen la formación de los compuestos deverbales.

Tenemos que esperar hasta la tercera recolección para ver un verbo del grupo -IR correctamente flexionado, lo cual solo ocurre con tres participantes y en un total de cinco respuestas. Para dominar la forma de los compuestos deverbales, los no nativos necesitan, por un lado, dominar la flexión verbal en presente y, por otro, recibir suficiente *input* como para entender y asimilar el proceso de formación de los deverbales en español. Después de 200 horas de

enseñanza en español, podemos suponer que nuestros participantes dominan hasta cierto punto la flexión verbal en presente, por lo que los errores en la forma de los deverbales se podrían atribuir más bien a la ausencia de conocimientos formales sobre estas expresiones en español, probablemente por falta de *input*, ya que se trata de un tema que a menudo se deja de lado en la enseñanza de ELE (Serrano Dolader, 2018, p. 310).

3.8 Discusión

Una vez analizado por separado cada tipo de compuesto, en esta sección consideramos las dos estructuras de manera conjunta para contrastar los resultados con nuestras hipótesis y tener una visión más global de la producción de compuestos por parte de nuestros participantes.

3.8.1 Las etapas iniciales en español L3

Conforme a los modelos de transferencia de adquisición de L3 que consideramos en esta tesis, era posible postular una influencia o bien a partir del euskera, o del francés. Según el *Typological Primacy Model* (TPM), el *Linguistic Proximity Model* (LPM) y el *Cumulative Enhancement Model* (CEM), la fuente de influencia debería ser el francés. En cuanto al *L2 Status Factor*, la influencia debería proceder del francés solo si se trata de la L2 de los participantes. En el caso contrario, se debería observar una influencia del euskera. Con respecto al llamado *L1 Factor*, la influencia debería proceder del francés solo si se trata de la L1 de los participantes y del euskera en el caso contrario. Finalmente, según el *Competition Model* (CM), la influencia principal debería proceder de la lengua dominante, una predicción similar al *L1 factor*. Sin embargo, considerando los datos de la prueba de control en euskera, resulta posible que la estructura N+N en la lengua-I de nuestros participantes no corresponda en todos puntos a lo que se comenta en la literatura, por lo que no queda claro en qué medida puede influir la estructura N+N del euskera en los N+N del español L3. Para los compuestos deverbales, es posible que las diferencias obvias entre la forma de los compuestos deverbales en euskera y los del español lleven a los aprendientes a un proceso de diferenciación que podría provocar más bien una influencia del francés. En ese sentido, una influencia exclusiva del francés para todos los participantes podría ser compatible con el CM.

Considerando estas hipótesis de influencia, tuvimos que detallar los posibles efectos de cada lengua sobre el español. A lo largo de esta sección, reproducimos las diferentes predicciones para

los compuestos N+N (N1-5) y para los deverbales (D1-6), las cuales nos servirán de punto de partida para orientar la discusión.

En cuanto a la productividad, habíamos planteado las siguientes posibilidades:

- N1-FR: Todos los participantes que muestren cierta influencia del francés deberían producir más compuestos sintagmáticos preposicionales como “un barco de pirata”, descartando en muchos casos la opción del compuesto N+N como “barco pirata”.
- N1-EU: Todos los participantes que muestren cierta influencia del euskera deberían producir una mayoría de compuestos N+N y evitar los compuestos preposicionales u otras formas alternativas.
- D1-FR: Los participantes que muestren una influencia del francés deberían producir un gran número de compuestos sintagmáticos preposicionales como “limpiador de botas” en lugar de compuestos deverbales propios como “limpiabotas”.
- D1-EU: Los participantes que muestren una influencia del euskera deberían producir una mayoría de compuestos deverbales sin nexo, como “limpiabotas”, “limpiador botas”, “botas limpiador” o “botas limpia” y evitar los compuestos sintagmáticos preposicionales como “limpiador de botas”.

En cuanto a los compuestos N+N, la tasa de producción de los grupos experimentales se diferencia significativamente de la del grupo de control, pero el que más se acerca es el grupo EU. Sin embargo, los grupos experimentales no se diferencian de manera significativa entre sí y, a pesar de su diferencia con el grupo de control, su tasa de producción de N+N es de un 75,3 %, lo que muestra que los participantes son capaces de producir compuestos N+N desde sus etapas iniciales en español, que es algo que hacen también en francés (60.91 %) y en euskera (77.64 %). En cuanto a la tasa más baja de compuestos N+N en la prueba de control en francés, consideramos que algunos ítems eran menos propensos a la producción de compuestos N+N en francés que en español, ya que la tarea experimental fue diseñada con vistas a elicitar compuestos N+N en español y la versión francesa de la tarea consiste en una adaptación con modificaciones mínimas de esta tarea. Si bien nos esforzamos por elegir ítems compatibles en las tres lenguas del estudio, es posible que algunos de ellos no sean completamente comparables, por lo que estas diferencias de productividad de una lengua a otra pueden estar vinculadas a los ítems seleccionados, más que a los procesos morfológicos más generales. A pesar de ello, los ítems en francés y en euskera nos permitieron observar una cierta productividad de las estructuras esperadas, sin que hubiera diferencias significativas entre los grupos experimentales en sus dos primeras lenguas.

Con respecto a los compuestos deverbales, los grupos experimentales produjeron un 75.00 % de sus respuestas con esta forma, sin diferencias significativas entre los grupos

experimentales, pero sí con una diferencia entre estos y el grupo de control (87,50 %). A pesar de las diferencias entre el grupo de control y los grupos experimentales, esos resultados nos muestran que los participantes son capaces de producir compuestos deverbales desde sus etapas iniciales, algo que hacen también en euskera (76,80 %) y, en menor medida, en francés (30,05 %). Una vez más, atribuimos esta tasa más baja de productividad a los propios ítems, que habían sido diseñados principalmente para elicitación de deverbales en español. Según estos resultados, el francés presenta más restricciones semánticas en la producción de los compuestos deverbales, por lo que, en gran medida, nuestros participantes produjeron compuestos sintagmáticos preposicionales con un verbo derivado (p. ej. “*un cireur de chaussures*”, es decir, “un lustrador de zapatos”) o formas simples derivadas a partir del verbo en lugar de las respuestas esperadas. Retomaremos la discusión sobre estos ítems más adelante al hablar del rasgo de animación.

Tanto en el caso de los N+N como de los deverbales, observamos que la tasa de producción de los compuestos en euskera se acerca más a la tasa del español, por lo que es posible que la productividad de los compuestos en euskera haya favorecido la producción de sus formas equivalentes en español, a pesar de que la estructura de los compuestos producidos en euskera es distinta de la del español, particularmente en cuanto a la direccionalidad. Además, como mencionamos antes, es posible que la tasa inferior de producción de compuestos en francés se deba a la confluencia de un efecto de la tarea y de la mayor restricción que presentan los compuestos del francés, algo que los participantes no parecen aplicar a su español L3.

Siguiendo este razonamiento, cuando tienen que acuñar una palabra en su L3 de nivel elemental, los participantes acuden a los mecanismos de composición y no parecen aplicar, por defecto, las restricciones semánticas de una de sus dos primeras lenguas, el francés. Esto, lo podríamos atribuir a dos posibilidades: 1) de manera general, a la hora de adquirir una nueva lengua, los mecanismos de composición son productivos para acuñar nuevas palabras o 2) el hecho de conocer una lengua que constriñe poco la creación de compuestos (el euskera en este caso) favorece la productividad del mecanismo de composición en la lengua no nativa. En cuanto a la primera posibilidad, estaríamos hablando de una estrategia que se ha observado en los niños a la hora de adquirir una L1 (Clark et al., 1985) y en la interlengua de los aprendientes de L2 (Liceras & Díaz, 2000b). En cuanto a la segunda posibilidad, solo podemos especular, ya que no tenemos datos con otros grupos de L2/L3 español de otras combinaciones de lenguas para comparar con los datos presentados aquí. Sin estos datos adicionales, resulta imposible determinar con fiabilidad si

lo observado se debe a una estrategia más general de la interlengua o si se puede atribuir a una influencia interlingüística. Entonces, concluimos con respecto a las predicciones N1 y D1 que la producción de nuestros participantes no se inclina de manera clara hacia ninguna de las dos fuentes de influencia posible.

En lo relativo a la direccionalidad, habíamos planteado las siguientes predicciones:

- N2-FR: Los participantes que muestren una influencia del francés deberían producir compuestos N+N con la direccionalidad adecuada, es decir, con el núcleo inicial.
- N2-EU: Los participantes que muestren una influencia del euskera deberían producir compuestos N+N con la direccionalidad inversa, es decir, con el núcleo final.
- D2-FR: Los participantes que muestren una influencia del francés deberían producir compuestos deverbales con la direccionalidad adecuada, es decir, el verbal primero, seguido por un sustantivo.
- D2-EU: Los participantes que muestren una influencia del euskera deberían producir compuestos deverbales con la direccionalidad inversa, es decir, un sustantivo seguido por un verbal.

En cuanto a los N+N, tras el análisis del euskera, habíamos planteado la posibilidad de que el francés influyera ya en la direccionalidad de los compuestos N+N en euskera. Efectivamente, los grupos experimentales produjeron una mayoría de compuestos N+N con el núcleo inicial en sus tres lenguas: francés (97,35 %), euskera (57,97 %) y español (88,90 %). En el caso del euskera, esta direccionalidad corresponde al orden inverso de lo que se documenta en la literatura, por lo que postulamos una posible influencia del francés sobre el euskera.¹⁸⁴ A su vez, esta influencia parece ayudar a los participantes a producir compuestos N+N con la direccionalidad adecuada en español.

Con respecto a los deverbales, la totalidad de las respuestas producidas por los grupos experimentales corresponden a la direccionalidad canónica del español, es decir la misma que en francés (V+N) y la inversa del euskera (N+V+sufijo). En este caso, los compuestos producidos por los participantes en sus dos primeras lenguas corresponden a la direccionalidad esperada, tanto en francés (98,60 %) como en euskera (99,22 %). Podemos confirmar con esos datos que la direccionalidad de los compuestos en euskera no influye de manera negativa en la direccionalidad de los compuestos en el español L3 de nuestros participantes.

¹⁸⁴ Profundizaremos este aspecto al volver específicamente sobre los resultados en las dos primeras lenguas en la sección 3.8.3.

Considerando a la vez la direccionalidad de los compuestos N+N y de los deverbales, sobresale la influencia del francés. De manera similar, en trabajos anteriores sobre adquisición de L2 (Liceras & Díaz, 2000b; Liceras & Valenzuela, 1998; Slabakova, 2002), se observó un efecto facilitador de direccionalidad de la L1 sobre la L2 cuando esta coincidía en ambas lenguas (p. ej., francés y español) y un efecto inhibitor cuando había conflicto de direccionalidad (p. ej., inglés y español). En nuestro caso, como los participantes tenían una lengua con el núcleo inicial (francés) y otra con el núcleo final (euskera), los dos efectos eran posibles, pero lo que observamos es una influencia facilitadora del francés para todos los grupos de participantes. Es decir, la que parece responsable de esta influencia facilitadora del francés en la direccionalidad de los compuestos en español L3 es la proximidad tipológica o estructural, no la edad ni el orden de adquisición de las dos primeras lenguas. Por su parte, el rasgo de animacidad parece modular el tipo de respuestas producidas. A este respecto, reproducimos las predicciones postuladas anteriormente:

- N3: Los participantes deberían producir una mayor tasa de compuestos N+N con los ítems cuyo referente es animado, ya que estos muestran un mayor grado de unidad de los significados asociados a cada lexema.
- D3-FR: Los participantes que muestren cierta influencia del francés deberían producir menos compuestos deverbales con los ítems animados, ya que los compuestos deverbales de esa lengua se refieren más a menudo a instrumentos que a agentes animados.
- D3-EU: El rasgo de animacidad no debería influir sobre la productividad de los deverbales para los participantes que muestren cierta influencia del euskera. Si se observa un cambio en función de este rasgo, el cambio debería afectar la forma y no la productividad como tal.

Los datos de nuestros grupos experimentales corroboran la predicción sobre los compuestos N+N en las tres lenguas, con una diferencia significativa en la productividad de los N+N animados vs. inanimados en euskera (92,30 % vs. 62,98 %), en francés (82,66 % vs. 42,53 %) y en español (87,36 % vs. 63,18 %). El grupo de control exhibe el mismo patrón también en español L1 con una tasa del 97,61 % de compuestos N+N en la condición [+animado] y del 73,81 % en la condición [-animado].

En cuanto a los deverbales, los datos recopilados en francés son llamativos en cuanto al rasgo de animacidad, que modula su producción de manera significativa. Concretamente, la tasa de producción de compuestos deverbales en francés es de apenas un 11,54 % en la condición [+animado] y asciende al 44,95 % en la condición [-animado]. En cambio, en euskera, la diferencia no es significativa y, además, se producen más compuestos deverbales en la condición [+animado]

(81,25 %) que en la condición [-animado] (72,75 %). El patrón del francés parece influir en la producción de los deverbales en español, ya que se observa un efecto significativo en la misma dirección por parte de todos los grupos no nativos, con una tasa de un 80,52 % deverbales en la condición [-animado] y un 69,47 % en la condición [+animado].¹⁸⁵

Entonces, encontramos un efecto contrario para los dos tipos de compuestos: el rasgo [+animado] favorece la producción de los N+N, mientras que el valor [-animado] favorece la producción de deverbales. En cuanto a los N+N, estos resultados no permiten concluir nada en relación con la influencia interlingüística, ya que parece ser una tendencia compartida por las tres lenguas. Sin embargo, nos permite reafirmar lo observado por Licerias y Klassen (2019) en relación con los compuestos del español. Como apuntan las autoras, el efecto podría proceder del hecho de que, cuando se trata de un compuesto de referente inanimado (p. ej., “carta bomba”), la percepción de separabilidad de los dos elementos es mayor que cuando se trata de un referente animado (p. ej., “hombre lobo”). Así, los compuestos animados muestran una mayor unidad entre sus componentes, mientras que los compuestos inanimados, al ser más separables, muestran una mayor flexibilidad frente a la inserción de afijos o preposiciones entre sus dos componentes. En lo referente a los deverbales, nuestros resultados son compatibles con los datos del estudio de corpus de Rosenberg (2008) sobre la productividad de los compuestos deverbales en francés y el estudio de Lacroix (en preparación), según los cuales los dos tipos de compuestos deverbales más productivos en francés son los instrumentales [-animado] primero, y los agentivos [+animado] en segundo lugar. En el estudio de Lacroix (en preparación), los deverbales [+animado] representan apenas el 20,04 % de los compuestos deverbales registrados en el diccionario. Esto parece indicar que nuestros participantes de los grupos experimentales producen, en español, unos compuestos deverbales que corresponden a los mismos criterios semánticos que en francés. Si bien se documenta el mismo efecto de animacidad en la producción de los deverbales en español, no lo hemos observado en nuestro grupo de control, probablemente porque todos los estímulos de la tarea fueron diseñados para favorecer la producción de compuestos deverbales. Considerando esto, nuestros resultados apuntan hacia la influencia del francés en la L3 de nuestros grupos experimentales en cuanto al criterio de animacidad de los compuestos deverbales, una influencia inicial que tiende a atenuarse con el tiempo, como veremos en la discusión sobre la evolución longitudinal de los participantes.

¹⁸⁵ Recordemos que los nativos de español produjeron casi un 100 % de deverbales en ambas condiciones.

En el estudio, nos interesamos también por el efecto potencial de los compuestos atestiguados, que nos llevó a establecer la siguiente hipótesis:

N4/D4: Al tratarse de unidades léxicas presentes en la lengua y, por lo tanto, susceptibles de formar parte del *input* recibido por los aprendientes, los ítems atestiguados deberían elicitar una mayor tasa de respuestas esperadas (N+N con núcleo inicial / V+N) para todos los participantes.

Nuestros datos no nos permiten confirmar esta hipótesis para nuestros grupos experimentales, ya que no observamos ningún efecto significativo del rasgo atestiguado para ambos tipos de compuestos ni sobre la productividad, ni sobre la direccionalidad. Es probable que los participantes no hayan recibido muestras suficientes de los compuestos en su *input*, por lo que no había ninguna diferencia entre los ítems atestiguados y no atestiguados para ellos.

El resto de las hipótesis de influencia inicial se relacionaban exclusivamente con los compuestos deverbales, ya que muestran una serie de particularidades en lo que respecta a la morfología tanto del verbo como del sustantivo que le acompaña.

D5-FR: Los participantes que muestren una influencia del francés deberían producir compuestos deverbales con un elemento verbal flexionado correctamente (en tercera persona del singular). Sin embargo, por lo menos en la primera recolección, es posible que aparezcan formas idiosincrásicas para el elemento verbal, es decir, o el verbo en infinitivo como “limpiarbotas”, o una flexión verbal alternativa como “limpiobotas” (sobre todo la primera persona del singular según Salomaa-Robertson, 2000).

D5-EU: Los participantes que muestren una influencia del euskera deberían derivar el verbal a partir de un proceso similar a los deverbales del euskera, es decir, con un sufijo agentivo o instrumental que podría corresponder a -dor(a) en español.

D6: Los participantes que muestren una influencia de cualquiera de sus dos primeras lenguas deberían, por lo menos al principio, producir los compuestos deverbales sin la -s final.

En los datos de español L3, encontramos muy pocos elementos deverbales derivados mediante un sufijo agentivo o instrumental, lo que excluye la influencia del euskera. En cambio, encontramos varias formas idiosincrásicas en los elementos verbales, lo que apunta hacia una posible influencia del francés neutralizada por el bajo nivel de dominio que los participantes tienen de la flexión verbal en español. Dicho de otro modo, consideramos estas formas idiosincrásicas como intentos de reproducción de la forma del francés en español, pero dado que los participantes todavía no dominan la flexión verbal del español, el resultado no corresponde a las respuestas esperadas.

En cuanto a la -s final, esta parece representar dificultades para todos los participantes, como planteábamos en nuestra hipótesis, lo cual había sido observado en estudios anteriores de adquisición de L2. Este último aspecto no apunta necesariamente hacia una influencia interlingüística, sino que pone de relieve la dificultad que representa este rasgo en los compuestos deverbales del español.

Para explicar las causas y el origen posible de las numerosas idiosincrasias en la forma de los compuestos deverbales, argumentamos que, a falta de un mayor dominio de la flexión verbal y de las características morfológicas de este tipo de compuestos, nuestros participantes informaban su producción a partir de indicios provenientes de los propios ítems. Concretamente, cuando el estímulo incluía una -s, los participantes reproducían una -s; cuando incluía un verbo conjugado en primera persona, reproducían una forma en primera persona; cuando incluía un infinitivo, reproducían una forma parecida al infinitivo (con o sin la -r final del verbo en infinitivo).

En vista de todo lo anterior, podemos afirmar que, de manera general, los tres grupos experimentales se comportan de forma comparable y que, cuando se observa una influencia interlingüística, se trata de un efecto facilitador del francés, independientemente de la edad o el orden de adquisición de dicha lengua. Estos resultados son compatibles con las predicciones del TPM, del LPM y del CEM, que asumen que la influencia interlingüística procede mayoritariamente de la lengua tipológica o estructuralmente más cercana, en nuestro caso, el francés. En cambio, las predicciones que asumen un efecto del orden de adquisición, como el *L2 Status Factor* o el *L1 Factor*, no son compatibles, ya que, según el perfil de los participantes, el francés puede corresponder tanto a la L1 como a la L2. En lo que respecta al CM, como mencionamos al inicio de la sección 3.8.1, los resultados pueden ser compatibles con este modelo si consideramos que la influencia del euskera, aun cuando es lengua dominante, no se aplica dada la divergencia entre esta lengua y el español con respecto a las estructuras estudiadas.

Si bien las predicciones de estos modelos nos permiten arrojar luz sobre los efectos más categóricos observados en las características morfosintácticas de los compuestos producidos, las perspectivas basadas en el uso nos permiten examinar las restricciones semánticas de los compuestos en una lengua no nativa, que es un tema que ha recibido menos atención. Recurrir a las nociones de construcción prototípica y de frecuencia nos permite explicar, entre otras cosas, por qué los participantes tienden a reproducir en español compuestos deverbales que muestran

características formales y semánticas semejantes a los deverbales en francés. Considerando que, en una construcción, el significado es indisociable de la forma, no resulta sorprendente que, al encontrarse con una nueva forma (en español) similar a otra que conocen (en francés), los participantes apliquen las mismas restricciones de significado. En nuestros datos, podemos observar esta tendencia en la preferencia que muestran los participantes de L3 por aplicar la estructura [V+N] a los referentes inanimados y la estructura [V-*eur*+de+N] a los animados. Hemos visto que, si bien los compuestos deverbales animados no son imposibles en francés, son mucho menos frecuentes que los inanimados y esto parece influenciar la producción de los deverbales en español L3.

3.8.2 Evolución longitudinal en español L3

Además de las hipótesis de influencia inicial, nos interesaba estudiar también el desarrollo de los mecanismos de composición a lo largo del proceso de adquisición:

- N5: Dado que todos los estímulos de nuestra tarea sirven para elicitare compuestos N+N bien formados en español (atestiguados o no), la tasa de producción de N+N no debería reducirse a pesar de la menor productividad de esta estructura en español, pero la tasa de errores en cuanto a la direccionalidad sí debería disminuir.
- D7: Con una mayor experiencia y un mayor *input*, todos los participantes deberían converger hacia un mayor porcentaje de aciertos en cuanto a la direccionalidad. El uso de formas verbales idiosincrásicas debería decaer con el tiempo y una mayor exposición a la lengua, cuando los aprendientes vayan dominando la flexión verbal en presente del indicativo. Sin embargo, en cuanto a la flexión verbal y el uso de la -s final, los estudios anteriores en L2 nos muestran que los errores persisten hasta niveles avanzados, por lo que la tasa de error debería decaer de manera menos pronunciada y representar un mayor desafío que la direccionalidad.

Tanto para los N+N como para los deverbales, nuestros datos nos han permitido comprobar que, cuando existían problemas de productividad y de direccionalidad en la etapa inicial, estas dificultades se atenuaban a medida que los participantes progresaban en su aprendizaje del español. En cambio, los errores iniciales vinculados a la forma (flexión verbal y -s genérica) de los compuestos deverbales perduran en el tiempo. En cuanto a la -s genérica, no se observaron diferencias significativas entre los grupos a lo largo de las diferentes recolecciones, sin embargo, los datos de la tercera recolección en cuanto a la flexión verbal son de particular interés. Efectivamente, en esta última etapa, el grupo EU se distingue de los demás en que muestra un mayor dominio de la flexión verbal. Como posible explicación, consideramos que este grupo, al

tener el euskera como lengua dominante, es más sensible a la morfología, lo que facilita su producción de verbos flexionados correctamente. No consideramos este fenómeno como una influencia interlingüística en sí, sino más bien como el resultado de una consciencia metalingüística más desarrollada. En la misma línea, recordamos al lector que, en el capítulo anterior, también habíamos observado un efecto distinto de la morfología en el grupo EU. Concretamente, el grupo EU usaba más la terminación (transparente u opaca) de los sustantivos para asignar el género a los sustantivos del español.

3.8.3 Los compuestos en las dos primeras lenguas (euskera y francés)

Para concluir esta discusión, queremos comentar dos elementos que surgieron en nuestros resultados y que, si bien no formaban parte de los objetivos de nuestro estudio de manera directa, representan una aportación considerable sobre el uso de las dos primeras lenguas de nuestros participantes. Las producciones de nuestros participantes en euskera presentaron diferencias con respecto a lo esperado en función de la literatura consultada sobre el tema de los compuestos, tanto para los N+N como para los deverbales y también observamos algunas particularidades en lo referente a los compuestos deverbales animados en francés.

3.8.3.1 Los N+N y los deverbales en euskera

Con respecto a los N+N, se considera generalmente que esos compuestos en euskera tienen un núcleo final, salvo en algunos casos excepcionales (Coyos, 2004, p. 58). Según lo que encontramos en nuestros datos, nuestros participantes mostraron mucha variación en cuanto al orden de los componentes, con un 59,62 % de compuestos con el núcleo inicial, el orden opuesto a lo esperado. Postulamos dos posibilidades para explicar este fenómeno: 1) o se trata de una influencia del francés o 2) el euskera de nuestros participantes se aleja de lo que plantean los trabajos descriptivos con respecto a la producción de compuestos N+N. La influencia a partir del francés representa una posibilidad, ya que la direccionalidad observada en varias respuestas N+N del euskera coincide con la del francés.

A su vez, la segunda posibilidad tiene sus fundamentos en los casos considerados como inhabituales en euskera, es decir, los casos que se presentan cuando el no-núcleo cumple una función similar a la del adjetivo. Dado que los compuestos N+A se forman con el núcleo a la izquierda, como vimos en la sección 3.2.3.1, esta direccionalidad corresponde a la esperada. Como

presentamos en la sección 3.2.1.1 sobre los compuestos N+N en español, el hecho de que un modificador esté considerado como un sustantivo adjetivado en lugar de un sustantivo puede cambiar la flexión de número del elemento no-nuclear del compuesto. En el ejemplo (3.174), si se considera que “clave” es un sustantivo, no llevará la marca de plural (a), pero si se considera que es un sustantivo adjetivado, la llevará (b).

- (3.174) a. los problemas clave
b. los problemas claves

La gramática del español recoge esta posible alternancia que solo afecta a la flexión de número (RAE & ASALE, 2010a, p. 193). En cambio, en euskera, dentro de un compuesto, el hecho de que a una palabra se le atribuya un valor adjetival puede cambiar su orden de aparición. Al interpretar nuestros resultados en euskera, constatamos que la mayoría de nuestros ítems se podían interpretar como compuestos atributivos, una interpretación que Coyos (2004) asimila a los compuestos N+A y, por lo tanto, a los compuestos con núcleo inicial. Entonces, es posible que los ítems seleccionados en nuestra tarea experimental hayan producido compuestos que son, en el uso cotidiano, poco habituales y con la direccionalidad contraria a los demás compuestos N+N del euskera. Sin embargo, a pesar de esta explicación, que se ajusta a una parte de los datos, tenemos otros ítems que no podemos considerar compuestos atributivos, como “coche bomba” y “carta bomba” que elicitaban compuestos con direccionalidad variable en euskera. A pesar de ello, recordemos que estos ejemplos muestran también divergencias a nivel interlingüístico, ya que, en español, este referente se interpreta como un tipo de carta, dando lugar al compuesto “carta bomba”, mientras que, en inglés, por ejemplo, se interpreta como un tipo de bomba, dando lugar al compuesto “*letter bomb*”.

En cuanto a los deverbales, la divergencia con respecto a lo esperado no procede de la direccionalidad de los compuestos, sino de los sufijos que se utilizaron para formar los deverbales. Si bien en la literatura se considera que los deverbales con valor agentivo se derivan mediante el sufijo *-tzaile* y que los instrumentales se derivan mediante el sufijo *-gailu* (Azkarate, 1995), nuestros datos revelan un patrón de uso diferente. El sufijo *-tzaile* es omnipresente tanto con los agentivos como con los instrumentales, mientras que el sufijo *-gailu* es reemplazado en varias respuestas por el sufijo *-tzeko*. Recordamos que el sufijo *-tzeko* es ante todo considerado por los gramáticos como la combinación del nominalizador de verbo *-tze* y del adnominal *-ko*, que produce

un significado que podríamos parafrasear como “que sirve para ...”. Los trabajos descriptivos consultados que hablan del sufijo *-tzeko* muestran sobre todo ejemplos en los que *-tzeko* encabeza una cláusula adnominal, como en (3.175), que tomamos de Oyharçabal (1998, p. 41).¹⁸⁶

(3.175) *Inglesa ikasteko ikasbide berria erosi dut*
 Inglés.DET.SG aprender.TZEKO método nuevo.DET.SG comprar AUX
 “He comprado un nuevo método para aprender inglés”

A su vez, uno de los ejemplos de Artigoitia (1991, p. 708), que reproducimos en (3.176), se acerca más a los compuestos deverbales producidos por nuestros participantes.

(3.176) *Lertxundik tristeak kontsolatzeko makina asmatu zuen*
 Lertxundi.ERG triste.DET.PL consolar.TZEKO maquina inventar AUX.PAST
 “Lertxundi inventó una maquina para consolar a las personas tristes”

Notamos de paso que “tristeak kontsolatzeko makina” tiene un significado que podríamos expresar inventando un compuesto deverbal en español: “(maquina) consolatristes”. Observamos también que “triste” no aparece en su forma escueta como suele ser el caso con los compuestos deverbales terminados en *-tzaile* y *-gailu* sino que aparece sufijado con un determinante plural.

Con una interpretación similar, el sufijo *-tzeko* también aparece en algunos deverbales lexicalizados, como “*irekitzeko*” (a partir del verbo “*ireki*”, “abrir”) que significa “abridor”, es decir, algo que sirve para abrir. Resulta curioso que, en su trabajo exhaustivo sobre los compuestos deverbales, Azkarate (1995) ni siquiera mencione *-tzeko* como un sufijo formador de deverbales. Es posible que se trate de un uso restringido a la zona dialectal de nuestro estudio, pero en nuestros datos este sufijo aparece a menudo. Reproducimos en (3.177) algunos ejemplos de las producciones de nuestros participantes.

(3.177) a. *kable mozteko*
 cable cortador
 b. *kontserbak irekitzeko*
 lata.DET-PL abridor
 c. *azala kentzeko*
 piel.DET-SG quitador
 d. *patata zuritzeko*
 patata pelador

¹⁸⁶ Oyharçabal la llama cláusula adnominal, Artigoitia la considera como un tipo de infinitiva relativa, más precisamente, una cláusula de propósito (p. 701)

d.	azietak	garbitzeko
	traste.DET-PL	limpiador
e.	botoila	hausteko
	botella.DET-SG	rompedor
f.	melo	jateko
	melón	comedor
g.	bixkotx	botatzeko
	pastel	lanzador
h.	munstroa	harrapatzeko
	monstruo.DET-SG	atrapador
i.	hortzak	ateratzeko
	diente.DET-PL	sacador

En estos ejemplos, algunas producciones aparecen con un complemento escueto, otras con un determinante singular o plural. Es posible que este sufijo no aparezca en las obras consultadas por esta razón precisa, es decir, no se considera que forme compuestos deverbales propios, sino que simplemente forma sintagmas con un verbo derivado. Sea cual sea la razón de la ausencia del sufijo *-tzeko* en la documentación consultada sobre los compuestos, nos parece relevante destacar esta producción de nuestros participantes para que se tome en cuenta en futuros trabajos.

Tanto para los compuestos N+N como para los deverbales en euskera, queremos reiterar que no hemos encontrado trabajos de adquisición de L1 sobre los compuestos, por lo que hemos tenido que guiarnos únicamente por los tratados descriptivos. Considerando este vacío en la literatura consultada, proponemos dos pistas de reflexión. Por una parte, es posible que la formación de los compuestos en el dialecto hablado por nuestros participantes (navarro-labortano y, en algunos casos, suletino) no concuerde en todos los aspectos con la lengua estándar (*euskera batua*) que se describe en las gramáticas del euskera (Berro Urrizelki et al., 2019; de Rijk, 2008; Hualde & Ortiz de Urbina, 2003; Laka, 1996) y en los trabajos específicos sobre los compuestos (Azkarate, 1993, 1995). Por otra parte, es posible que la lengua esté cambiando y que en nuestros datos estemos observando cambios que no se hayan documentado todavía.

3.8.3.2 Los deverbales en francés y el rasgo de animacidad

Con respecto a los deverbales animados en francés, observamos en nuestros datos una tasa bastante baja de compuestos deverbales (28,25 %), con una gran diferencia en función del criterio de animacidad, es decir, una tasa de un 11,54 % de compuestos deverbales en la condición [+animado] y un 44,95 % en la condición [-animado].

Consideramos que, de manera global, esta divergencia se debe esencialmente a los ítems de la tarea, ya que, como hemos mencionado, fueron diseñados originalmente para elicitación de compuestos deverbales en español. Especulamos que, para un ítem como “limpiabotas”, hubiéramos podido obtener más compuestos deverbales con una imagen en la que una persona porta una bandera ya que, en francés, existen algunas entradas de referentes animados con el equivalente del verbo “portar” (“*porte-étandard*”, “*porte-drapeau*” y “*porte-bannière*”, “*porte-fanion*”, etc.) y ninguna con el equivalente del verbo “limpiar/lavar” (“*laver*”). Es decir, en francés, el verbo “*laver*” solo se emplea para formar deverbales inanimados (“*lave-vaisselle*”, “*lave-auto*”, “*lave-glace*”, “*lave-main*”, *lave-linge*), mientras que en español, si bien se emplea sobre todo para los inanimados (“limpiabarros”, “limpiadientes”, “limpiaparabrisas”, “limpiaplumas”, “lavavajillas”, “lavamanos”, “lavadientes”, etc.), también existen deverbales que tienen acepciones principalmente animadas (“lavacoches”, “limpiachimeneas”) u otras que pueden referirse tanto a animados como a inanimados (“limpiacristales”). Sin embargo, más que un caso aislado, consideramos que este ítem es representativo de la menor productividad de los deverbales en francés para los referentes animados. Para más detalles sobre esta pista, referimos al lector al análisis de Lacroix (en preparación), pero mencionamos aquí la lista relativamente breve de verbos en francés encontrados en los compuestos deverbales animados. En (3.178), presentamos la lista para los animados no humanos (plantas, animales, monstruos) y en (3.179) la lista para los humanos.¹⁸⁷

(3.178) *attraper, casser, couvrir, croquer, garder, gober, gratter, hocher, manger, passer, percer, pincer, piquer, porter, pousser, tourner, tuer;*

(3.179) *aider, boucher, bouler, briser, casser, couper, crever, croquer, gagner, garder, gober, gratter, gripper, lècher, manger, pêter, pisser, porter, prêter, rabattre, risquer, serrer, songer, souffrir, trainer, troubler.*

En cambio, en español, la lista de verbos que se encuentran en los compuestos deverbales animados es más larga y, por cuestiones de concisión, no la reproducimos aquí y referimos el lector a Lacroix (en preparación). Con esto, queremos destacar que, como hemos mencionado anteriormente, el francés parece mostrar mayores restricciones en la productividad de los

¹⁸⁷ Otro dato importante sobre la productividad de los animados en el uso actual: el compuesto verbal animado más reciente registrado en Usito aparece en 1948 (*casse-pieds*), mientras que, desde este año, se han atestiguado más de cincuenta deverbales inanimados en el mismo diccionario (Lacroix, en preparación).

compuestos deverbales. Sin embargo, estas restricciones corresponden a grandes rasgos a las características enumeradas por Moyna (2011, pp. 202-203) sobre los compuestos deverbales animados en español. Los compuestos deverbales humanos suelen ser usados para denominar oficios poco prestigiosos (“aparcacoches” en español, *croque-mort* en francés para referirse a un enterrador) o, por lo menos vistos de manera negativa (“cagatinta” en español, *gratte-papier* en francés para referirse a un empleado de oficina). También se usan para aludir a un comportamiento negativo (“rompehuevos” en español, *cassepieds* en francés) o para exagerar el comportamiento de una persona (“chupamedias” en español, *lèche-botte* en francés). En cuanto a los compuestos deverbales animados no humanos, se suele aludir a una característica saliente de la planta (“quebrantapiedras”, *perce-pierre*) o del animal (“cascanuez”, *cassenoix*). Considerando esto, la diferencia no parece encontrarse en las características semánticas de los compuestos deverbales animados, ya que las dos lenguas muestran patrones similares, sino que el principal responsable de las diferencias entre las dos lenguas podría ser la frecuencia en el uso.

4 Discusión general y conclusiones

En el presente capítulo sintetizamos los tres primeros capítulos con el fin de articular una discusión general sobre la influencia interlingüística observada en nuestros datos. Para ello, resumiremos en primer lugar los diferentes modelos de adquisición presentados en el capítulo 1 y propondremos una síntesis de los resultados del segundo y del tercer capítulo. Esta primera parte nos servirá de base para la posterior discusión, la cual se orientará hacia la influencia interlingüística y las habilidades metalingüísticas. Posteriormente, presentaremos una breve reflexión sobre las posibles implicaciones de nuestros resultados para la enseñanza del español como L3. A modo de conclusión de la tesis, terminaremos el capítulo con una presentación de las principales aportaciones de nuestro trabajo y propondremos líneas de investigación para futuros trabajos.

4.1 Síntesis general

4.1.1 Modelos de adquisición de L3

Como hemos presentado en el capítulo 1, a lo largo de las dos últimas décadas, han surgido varias propuestas para modelizar las posibilidades de influencia interlingüística en el proceso de adquisición de la L3. Dichas propuestas se diferencian principalmente en cuanto al factor principal que condiciona la influencia: el orden de adquisición o la proximidad de las lenguas (a nivel global o más específico). Por otra parte, podemos clasificar los modelos en función de sus predicciones sobre el impacto de las lenguas previamente adquiridas en el proceso de adquisición de una nueva lengua, es decir, un papel facilitador, neutro o inhibidor. Finalmente, ciertos modelos predicen una influencia total, en línea con la hipótesis de la transferencia completa y del acceso completo de L2 (Schwartz & Sprouse, 1996), mientras que otros predicen una influencia propiedad por propiedad.

En cuanto al orden de adquisición, la propuesta de la transferencia exclusiva a partir de la L1 es generalmente atribuida a Hermas (2010, 2014) y Na Ranong y Leung (2009). Esta propuesta predice siempre una influencia a partir de la L1, ya sea esta facilitadora o inhibidora. De manera similar, pero esta vez a partir de la L2, la hipótesis del *L2 Status Factor*, inicialmente presentada por Williams y Hammarberg (1998) y luego profundizada por Bardel y Falk (2007), plantea la influencia a partir de la L2 como la única posibilidad, la cual, como en el caso anterior, puede ser

también facilitadora o inhibidora. Según esta hipótesis, la L2 actúa como un filtro que bloquea el acceso a la L1.

En cuanto a la proximidad tipológica entre las lenguas, el *Typological Primacy Model* (TPM) introducido por Rothman (2011) predice una influencia exclusiva (transferencia total) a partir de la lengua (L1 o L2) percibida como más cercana a la L3. Pese al nombre del modelo, la distancia entre las lenguas se considera en términos de psicotipología (Kellerman, 1983) y no de tipología, es decir, se considera la distancia como la percibe el hablante (de manera inconsciente) y no como la perciben los lingüistas. Dado el carácter total de la transferencia que opera según este modelo, la predicción es que la transferencia será inhibidora cuando las estructuras no coinciden entre la lengua fuente y la lengua meta y facilitadora cuando las dos lenguas confluyen. Por su parte, el *Cumulative Enhancement Model* (CEM) propuesto por Flynn, Foley y Vinnitskaya (2004) propone una influencia que puede proceder de la L1, de la L2 o de ambas, siempre y cuando esta sea facilitadora. Para evitar la redundancia que representaría la adquisición de una misma estructura en una lengua previamente adquirida y en la L3, las dos lenguas previamente adquiridas pueden contribuir al proceso de adquisición de la L3 a medida que el hablante va notando las semejanzas entre una lengua fuente y la L3. La propuesta más reciente, el *Linguistic Proximity Model* de Westergaard et al. (2017) se construye sobre las bases del TPM y del CEM, pero se distingue del TPM principalmente por el tipo de proximidad considerada y el abarque de la influencia, y del CEM porque predice tanto efectos facilitadores como inhibidores. El LPM se basa en la proximidad estructural para predecir la fuente de influencia interlingüística, es decir que, a pesar de que dos lenguas no se parezcan a nivel global, podría darse una influencia a nivel de una estructura precisa si hay semejanza en la propia estructura. Dado el tipo de proximidad considerado, este modelo supone una influencia a nivel de las propiedades y no una transferencia total, como predice el TPM. Si bien la condición necesaria para que la influencia opere es la semejanza estructural, este modelo predice una influencia inhibidora cuando la semejanza percibida en la superficie no corresponde completamente a la estructura de la lengua meta. Por último, el Modelo de competición (*Competition Model*, CM; E. Bates & MacWhinney, 1987; MacWhinney, 1987, 1997, 2008, 2012; entre otros) se basa en la percepción de indicios (*cues*) para establecer conexiones entre la forma y la función a largo del proceso de adquisición. En el proceso de adquisición de la L3, el CM predice que el aprendiente usará los indicios de su lengua dominante (por lo general, la L1) para procesar la L3, al tener esta un mayor nivel de activación. Sin embargo, este modelo plantea la posibilidad

de una influencia acumulativa y dinámica a lo largo del proceso de la L3, por lo que no excluye una posterior influencia de una lengua no dominante (por lo general, la L2).

4.1.2 El género gramatical

En el capítulo 2, presentamos una descripción del género gramatical en las tres lenguas del estudio y resaltamos las semejanzas entre el francés y el español, así como el contraste entre el español y el euskera. Además, vimos que el género gramatical representa un desafío para los aprendientes de español como lengua no nativa. A este respecto, ciertos factores que parecen afectar el dominio del género gramatical figuran de manera recurrente en los estudios anteriores sobre la adquisición del género gramatical en español. En concreto, observamos que los aprendientes tienden a sobreutilizar el género masculino en la concordancia con los sustantivos femeninos, pero no al revés, por lo que se considera generalmente que los aprendientes aplican el género masculino por defecto. El tipo de concordancia también tiene un impacto sobre la concordancia de género, en el sentido de que la concordancia con un determinante parece provocar menos errores que la concordancia con un adjetivo. La terminación de los sustantivos, es decir, la morfofonología, también afecta a la concordancia de género, ya que las terminaciones transparentes facilitan la concordancia, mientras que las terminaciones opacas la inhiben. Finalmente, el criterio analógico, es decir, la congruencia de género entre el sustantivo en español y el género del sustantivo equivalente en una de las lenguas previamente adquiridas también parece jugar un papel facilitador, ya que la incongruencia de género inhibe la concordancia en español.

De manera general, nuestros resultados se encuentran en línea con los resultados de las investigaciones anteriores sobre el género gramatical en español, ya que observamos más aciertos con el género masculino (masculino por defecto), con los determinantes (tipo de concordancia), con las terminaciones transparentes (morfofonología) y con sustantivos con género equivalente en español y en francés (criterio analógico). Ahora bien, los resultados de mayor importancia para la presente discusión están relacionados con el papel de la morfofonología y del criterio analógico, que parecen estar modulados por la lengua dominante de los participantes.¹⁸⁸ En particular, el grupo

¹⁸⁸ Recordamos al lector que, para clasificar a los participantes en función de su lengua dominante, usamos como criterio principal la(s) lengua(s) hablada(s) en casa. Esto nos dio tres grupos: solo euskera (grupo EU), solo francés (grupo FR) o ambas lenguas (grupo FR/EU).

dominante de francés (grupo FR) parece recurrir más al criterio analógico con el francés a la hora de hacer la concordancia en español, mientras que el grupo dominante de euskera (grupo EU) parece basarse más en la terminación del sustantivo cuando esta es transparente. Por su parte, el grupo que habla tanto el euskera como el francés en casa (grupo FR/EU) no muestra ningún patrón claro en función de los dos factores antes mencionados.

4.1.3 Los compuestos

En el capítulo 3, presentamos una descripción detallada de los mecanismos de composición de los N+N y deverbales en las tres lenguas del estudio y, de nuevo, subrayamos las semejanzas entre el español y el francés, por un lado, y el contraste entre el español y el euskera, por otro. Por otra parte, destacamos el hecho de que, en la adquisición del español como lengua no nativa, los estudios anteriores muestran que la direccionalidad (núcleo inicial o final) de los compuestos puede provocar dificultades en la formación de los compuestos N+N y deverbales. Además, en el caso concreto de los deverbales, la morfología verbal y de número parece presentar un desafío aún más importante para los hablantes no nativos del español. Por último, nos interesamos también en el rasgo de animacidad y en el hecho de que un compuesto esté atestiguado o no.

Nuestros resultados muestran que los participantes son capaces de producir los dos tipos de compuestos desde las etapas iniciales de su aprendizaje. De manera general, no observamos grandes dificultades en cuanto a la direccionalidad, lo cual parece indicar un efecto facilitador del francés, que muestra características semejantes al español. Las pequeñas diferencias observadas en cuanto al rasgo de animacidad para los deverbales en la primera recogida de datos disminuyen con el tiempo. Adicionalmente, no se observó ningún efecto causado por el hecho de que un compuesto esté atestiguado o no en ninguna de las recogidas y solo resultó ser un factor relevante para el grupo de control (hablantes nativos de español). En cambio, los errores vinculados a la forma (flexión verbal y -s genérica) de los compuestos deverbales son pronunciados en la primera recolección y perduran en el tiempo. De manera global, los tres grupos experimentales se comportan de forma comparable.

4.1.4 Comparación de los fenómenos

Antes de entrar en la discusión general, presentamos en la Tabla 4.1 un resumen de todas nuestras hipótesis relacionadas con una posible influencia interlingüística. Recordamos al lector

que el prefijo indica el fenómeno en cuestión, es decir, el género gramatical (G), la formación de los compuestos N+N (N) o los deverbales (D). La sigla de dos letras indica la lengua de influencia, es decir, el francés (FR) o el euskera (EU). Las dos últimas columnas corresponden a lo que hemos podido comprobar con nuestros resultados para los dos grupos que muestran los resultados más contrastados (EU y FR). El signo interrogativo (?) indica que nuestros datos no permiten comprobar ninguna hipótesis de manera clara. La marca de verificación (✓) indica que nuestros datos corroboran esta hipótesis para el grupo en cuestión y la equis (✗) indica que nuestros datos manifiestan evidencias en contra de la hipótesis para este grupo.

Tabla 4.1. Resumen de las hipótesis de influencia interlingüística.

	Hipótesis	Grupo EU	Grupo FR
G1-FR	Consideramos que los participantes que muestren una influencia del francés deberían beneficiarse de un efecto facilitador (mayor tasa de acierto de manera general), ya que los sistemas de género del español y del francés son similares.	?	?
G1-EU	El euskera no materializa el género gramatical, por lo tanto, una influencia de esta lengua podría manifestarse con una tasa de concordancia más baja y, por lo tanto, una sobreutilización del masculino, el género por defecto.	?	?
G2-FR	La influencia del francés debería modular el efecto del criterio analógico, es decir, el género analógico debería facilitar la concordancia de manera más acentuada en los participantes que muestren una influencia del francés.	✗	✓
G2-EU	Si bien predecimos un efecto facilitador de la terminación transparente para todos los grupos, predecimos un efecto de mayor amplitud para los participantes que muestren una influencia interlingüística a partir del euskera.	✓	✗
N1-FR	Todos los participantes que muestren una influencia del francés deberían producir más compuestos sintagmáticos preposicionales como “un barco de pirata”, descartando en muchos casos la opción del compuesto N+N, como “barco pirata”.	?	?
N1-EU	Todos los participantes que muestren una influencia del euskera deberían producir una mayoría de compuestos N+N y evitar los compuestos preposicionales u otras formas alternativas.	?	?
N2-FR	Los participantes que muestren una influencia del francés deberían producir compuestos N+N con la direccionalidad adecuada, es decir, con el núcleo inicial.	✓	✓
N2-EU	Los participantes que muestren una influencia del euskera deberían producir compuestos N+N con la direccionalidad inversa, es decir, con el núcleo final.	✗	✗

D2-FR	Los participantes que muestren una influencia del francés deberían producir compuestos deverbales con la direccionalidad adecuada, es decir, el deverbale primero, seguido por un sustantivo.	✓	✓
D2-EU	Los participantes que muestren una influencia del euskera deberían producir compuestos deverbales con la direccionalidad inversa, es decir, un sustantivo seguido por un deverbale.	✗	✗
D3-FR	Los participantes que muestren cierta influencia del francés deberían producir menos compuestos deverbales con los ítems animados.	✓	✓
D3-EU	El rasgo de animación no debería influir sobre la productividad de los deverbales para los participantes que muestren cierta influencia del euskera. Si se observa un cambio en función de este rasgo, el cambio debería afectar la forma y no a la productividad como tal.	✗	✗
D5-FR	Los participantes que muestren una influencia del francés deberían producir compuestos deverbales con un elemento verbal flexionado correctamente (en tercera persona del singular). Sin embargo, por lo menos en la primera recolección, es posible que aparezcan formas idiosincrásicas para el elemento deverbale; por ejemplo, puede aparecer el verbo en infinitivo, como “limpiarbotas”, o una flexión verbal alternativa, como “limpiobotas” (sobre todo la primera persona del singular, siguiendo a Salomaa-Robertson, 2000).	✓	✓
D5-EU	Los participantes que muestren una influencia del euskera deberían derivar el deverbale a partir de un proceso similar a los deverbales del euskera, es decir, con un sufijo agentivo o instrumental, que podría corresponder a -dor(a) en español.	✗	✗

De manera general, se puede observar en la Tabla 4.1 que la mayoría de las hipótesis de influencia a partir del francés han sido comprobadas, mientras que una sola hipótesis vinculada con el género (G2) ha sido corroborada para el euskera.¹⁸⁹

¹⁸⁹ Recordamos al lector que, con respecto al efecto de la terminación, el grupo EU obtuvo una tasa de aciertos del 70,92 % en la condición transparente frente a un 52,28 % en la condición opaca. A su vez, el grupo FR obtuvo una tasa de aciertos de 59,93 % en la condición transparente y de 52,29 % en la condición opaca. Por otra parte, el criterio analógico afectó de manera más considerable al grupo FR, con una tasa de aciertos del 71,6 % en la condición congruente y de 40,3 % en la condición incongruente. El grupo EU obtuvo una tasa de aciertos del 65,36 % en la condición congruente frente a un 59,06 % en la condición incongruente.

4.2 Influencia interlingüística

Recordemos que, al inicio del estudio, nuestros participantes de los tres grupos experimentales (FR, EU, FR/EU) tenían un nivel nativo o casi nativo en sus dos primeras lenguas, el francés y el euskera, con ciertas variaciones en el orden de adquisición y en el uso de estas. Por lo tanto, tenían un bagaje lingüístico con dos fuentes similares susceptibles de influir en el proceso de adquisición del español como L3. Considerando lo anterior, a continuación, repasaremos las implicaciones de las hipótesis de la Tabla 4.1 para los modelos de L3 antes presentados. La discusión se enfocará en los dos grupos que muestran las diferencias de mayor contraste, es decir, los grupos EU y FR.

4.2.1 El orden de adquisición

Dado que hemos observado una influencia similar del francés en la producción de los compuestos (N2-FR, D2-FR, D3-FR, D5-FR) para los dos grupos (EU, FR), a pesar del orden de adquisición opuesto de sus dos primeras lenguas, los modelos de AL3 basados en el orden de adquisición no parecen ser compatibles con nuestros resultados. Concretamente, la hipótesis de influencia exclusiva a partir de la L1 (Hermas, 2010, 2014; Na Ranong & Leung, 2009) predecía una influencia negativa del euskera en todos estos casos para el grupo EU, lo cual no concuerda con nuestros resultados. A su vez, la hipótesis de influencia exclusiva a partir de la L2 (Bardel & Falk, 2007) predecía el efecto contrario, es decir, una influencia negativa a partir del euskera para el grupo FR, lo cual tampoco hemos observado. Por otra parte, los resultados relacionados con el género gramatical (G2-FR, G2-EU) son menos claros en cuanto al papel del orden de adquisición de las lenguas. Para el grupo FR, la L1 parece tener un papel importante a la hora de aplicar el criterio analógico a los sustantivos del español, mientras que para el grupo EU, la terminación de los sustantivos parece anular este efecto cuando hay contradicciones entre los dos indicios. Si bien el criterio analógico solo se puede vincular a la influencia a partir del francés, el papel de la terminación es más ambiguo. En nuestra hipótesis G2-EU, vinculamos el mayor peso de la morfología a un efecto del euskera, siguiendo a Caffarra et al. (2017). Según estos autores, el hecho de tener como lengua dominante el euskera, una lengua aglutinante sin clasificación de género, favorece el procesamiento de los indicios formales y debilita las representaciones de género en el léxico mental. Por otra parte, también podríamos vincular la mayor dependencia de la forma a un efecto de la instrucción formal en francés. Falk et al. (2015) consideran que la instrucción formal

en la L2 favorece el desarrollo de conocimientos metalingüísticos y es precisamente uno de los factores que favorece la influencia interlingüística de la L2 sobre la L3. Dado que las dos lenguas suelen ser adquiridas en un mismo contexto de aprendizaje, los aprendientes suelen aplicar las mismas estrategias para ambas lenguas. Considerando esto, el efecto de recibir enseñanza formal sobre la concordancia de género en francés como L2 podría haber favorecido el desarrollo de estrategias metalingüísticas sobre los indicios de género en los sustantivos del francés y, posteriormente, del español. Recordemos que, si bien el francés presenta terminaciones menos transparentes que el español, existen ciertas terminaciones con un alto poder predictivo.¹⁹⁰ En ese sentido, es posible que los aprendientes que han adquirido el francés en un contexto más formal (grupo EU) hayan desarrollado una estrategia en esta lengua que luego aplican a las terminaciones del español. En cualquier caso, con respecto al género gramatical, resulta evidente que el grupo EU aplica una estrategia diferente a la del grupo FR, por lo que el orden de adquisición o la lengua dominante (por lo general, los dos coinciden en el caso de nuestros participantes) parece tener un impacto en el proceso de adquisición de la L3. Los resultados no son lo suficientemente claros como para ser compatibles ni con el *L1 factor*, ni con el *L2 status factor*, pero, al menos indican que el orden de adquisición y la lengua dominante son factores que tenemos que tomar en cuenta a la hora de estudiar la influencia interlingüística en AL3.

4.2.2 La proximidad tipológica o estructural

A su vez, el TPM (Rothman, 2011) tampoco parece enteramente compatible con nuestros resultados, ya que predice una transferencia exclusiva a partir de la lengua más cercana tipológicamente, en este caso, el francés. Para comprobar la predicción del TPM, hubiera sido necesario corroborar todas las hipótesis que presentan el francés como fuente de influencia o, al menos, no encontrar ninguna influencia a partir del euskera. Lo observado en cuanto a la formación de los compuestos (N2-FR, D2-FR, D3-FR, D5-FR) es compatible con las predicciones del TPM. Sin embargo, con respecto al género gramatical, el grupo EU presenta algo que se puede interpretar como una influencia a partir del euskera, es decir, la influencia de la morfología del euskera sobre

¹⁹⁰ Como mencionamos en el capítulo 2, ciertas terminaciones fonológicas como /ã/ (99 %), /ɛ̃/ (98 %), /o/ (93 %) y /ɛ/ (93 %) predicen el género masculino, mientras que /z/ (97 %) y /j/ (90 %) predicen el femenino. Sin embargo, para alcanzar un mayor poder predictivo para una mayor cantidad de sustantivos, es necesario recurrir a la terminación ortográfica (Lyster, 2006).

el procesamiento de las terminaciones de los sustantivos en español. Aun considerando que la mayor sensibilidad a las terminaciones de los sustantivos en español procede de un efecto de instrucción en francés (como hemos comentado en el párrafo anterior) y que, por lo tanto, solo se observa una influencia a partir del francés en todos los casos, el TPM no permite explicar por qué los dos grupos no muestran la misma sensibilidad a dichas terminaciones. Además, si la transferencia opera de manera completa, como predice el TPM, el criterio analógico debería aplicarse de la misma manera para ambos grupos, lo cual, como hemos visto, no sucede en nuestros datos.

Por su parte, el LPM (Westergaard et al., 2017) genera predicciones similares al TPM en cuanto a la proximidad entre las lenguas, pero contempla la posibilidad de una influencia a partir de más de una lengua previa, es decir, no presupone una transferencia exclusiva a partir del francés, como el TPM. Según el LPM, lo que guía el proceso de adquisición de una lengua son los micro-indicios (*micro-cues*) que se desarrollan a partir de la interacción entre la Gramática Universal (GU) y el *input*.¹⁹¹ La principal diferencia entre la adquisición de L1, de L2 y de L3 se debe a la acumulación de los micro-indicios existentes: para la adquisición de su L1, el hablante no tiene acceso a micro-indicios previos, para la adquisición de su L2, el aprendiente tiene acceso al inventario de micro-indicios de su L1 y, finalmente, para la adquisición de L3, el aprendiente tiene acceso al inventario de su L1 y de su L2. Los micro-indicios permiten el aprendizaje mediante análisis (*learning by parsing*) para cualquiera de las lenguas (L1/L2/L3) y, de manera más precisa, en el caso de la adquisición de la L3, permiten la comparación de las estructuras de las distintas lenguas y guían el proceso de influencia interlingüística. Siguiendo esta lógica, ya que las dos primeras lenguas de nuestros participantes comportan micro-indicios diferentes y las estructuras del francés son similares a las del español, podemos considerar que los micro-indicios del francés deberían ser los privilegiados por nuestros participantes a la hora de analizar el *input* en la L3. Es lo que parece ocurrir en el caso de la formación de los compuestos, ya que se observa una influencia exclusiva a partir del francés. Del mismo modo, la fuente privilegiada para la selección del género gramatical parece ser el francés, ya que todos los participantes realizan la concordancia de género

¹⁹¹ Si bien los micro-indicios son estructuras que se podrían parecer a las construcciones de los acercamientos basados en el uso, Westergaard enfatiza: “this approach also differs from a constructivist model in that the micro-cues are not surface strings of word combinations, but pieces of (hierarchical) I-language structure that make reference to linguistically relevant distinctions.” (Westergaard, 2009b, p. 1031)

en cierta medida. Sin embargo, otros micro-indicios parecen entrar también en juego a la hora de asignar un género a los sustantivos. En euskera, los morfemas de caso que aparecen al final de los sustantivos (y de otras categorías de palabras, como los adjetivos, los pronombres, etc.) pueden cumplir una variedad de funciones que se materializan de maneras diferentes en otras lenguas, como el orden de palabras o el uso de preposiciones.¹⁹² Por lo tanto, es posible que los aprendientes dominantes de euskera presten una mayor atención a estos micro-indicios alternativos en el momento de analizar su L3. Es posible que estos aprendientes analicen inicialmente las terminaciones más frecuentes de los sustantivos del español (-o, -a) según los micro-indicios de su lengua dominante (p. ej., la -a en “casa” podría corresponder a un artículo definido que sería el equivalente de “el/la” en español) y que, posteriormente, al darse cuenta de que este primer análisis es erróneo, reanalicen la función de la terminación y la asocien a una marca de género.¹⁹³ Esta interpretación del LPM y de los micro-indicios se distancia de la propuesta inicial de Westergaard, ya que el Modelo de los micro-indicios se aplica inicialmente a la sintaxis y las terminaciones de los sustantivos en español pertenecen más bien al ámbito del léxico y de la morfología. No obstante, dado que los morfemas de caso del euskera juegan un papel sintáctico, nos parece viable aplicar el modelo a la asignación del género en la combinación de lenguas estudiadas en el presente trabajo. Aparte de la proximidad estructural, Westergaard (2019) no descarta la posibilidad de que otros factores entren en juego a la hora de determinar la fuente de influencia, pero no queda claro cómo

¹⁹² A modo de ejemplo, al sustantivo “etxe” (casa), se le puede añadir diferentes sufijos para cumplir funciones que normalmente se cumplen con una o varias palabras independientes en lenguas no aglutinantes, como el español y el francés. Los ejemplos (i-iv) representan algunos casos posibles:

- (i) etxea
“la casa”
- (ii) etxean
“en la casa” (inesivo)
- (iii) etxera
“a la casa” (dirección, adlativo)
- (iv) etxetik
“de/desde la casa” (procedencia, ablativo)

¹⁹³ Con nuestros datos, resulta imposible discernir cuánto tiempo dura esta primera etapa de análisis erróneo y si dicha etapa está asociada, por ejemplo, a una mayor omisión de los determinantes en los sintagmas determinantes del español. Recordamos al lector que, si bien ciertos participantes omitieron los determinantes en la prueba de denominación de imágenes, no hemos encontrado ninguna relación entre la lengua dominante de los participantes y las tasas de omisión.

se integrarían estos otros factores en su modelo. Por ejemplo, nuestros datos revelan que no todos los grupos prestan la misma atención a los micro-indicios de la terminación de los sustantivos y que se trata de un efecto que parece interactuar con la lengua dominante. Por otra parte, resulta difícil interpretar lo que predeciría el LPM con respecto al criterio analógico, ya que este no se manifiesta como un micro-indicio, sino que se trata de un proceso de asociación interlingüística de los rasgos de género que afecta a las representaciones léxicas.

Si bien estos dos modelos que se basan en la proximidad, ya sea proximidad estructural (LPM) o tipológica (TPM), predicen de manera adecuada la influencia del francés en cuanto a la formación de los compuestos, no son del todo compatibles con nuestros datos sobre el género gramatical. Aunque la noción de micro-indicio del LPM proporciona un mayor nivel de precisión en las predicciones, nuestros datos revelan que parece imprescindible tomar en cuenta la lengua dominante como un factor determinante de la influencia interlingüística.

4.2.3 La lengua dominante

El CM también se basa en la noción de indicios (*cues*) para comparar las lenguas y determinar la fuente de influencia, lo que permite llegar a conclusiones similares al LPM acerca de la influencia del euskera en el análisis de las terminaciones en español L3. Adicionalmente, las predicciones del CM toman en cuenta el nivel de activación de las lenguas como un factor determinante de la influencia interlingüística. En el caso de nuestros participantes, consideramos que la lengua dominante es la lengua con el mayor nivel de activación y, por lo tanto, el CM predice una influencia a partir de la lengua dominante, es decir, el euskera para el grupo EU y el francés para el grupo FR, que son predicciones similares al *L1 factor*, como mencionamos anteriormente. Sin embargo, la hipótesis del *L1 factor* no explora la noción de activación como hace el CM. A modo de ejemplo, exploramos la noción de activación conforme a lo que hemos observado en cuanto al género gramatical con el sustantivo “cama”, que el grupo EU hace concordar en femenino en el 67 % de los casos, mientras que el grupo FR lo hace concordar en masculino en el 73 % de los casos. En lo que concierne al nivel de activación, podemos considerar que los participantes del grupo FR tienden a asociar las formas del español con las formas equivalentes en francés (“*lit*”) y, con ello, parecen extender a la L3 los rasgos que presentan estas unidades léxicas en su L1. En el caso de nuestro ejemplo, el género del francés correspondería al masculino, lo cual no representa la opción adecuada. En cambio, los participantes del grupo EU asocian las formas del español con

sus equivalentes en euskera, lo que no les proporciona un acceso directo al rasgo de género, ya que el euskera no materializa el género. Por lo tanto, tienen que buscar otra estrategia para asignar un género a los sustantivos del español, es decir, analizar la terminación de los sustantivos, como comentamos anteriormente. En cuanto a los compuestos N+N (hipótesis N2-FR), el CM también parece compatible con nuestros datos si tomamos en cuenta la baja productividad de los N+N con la direccionalidad esperada en euskera. La estructura N+N no parece ser utilizada con mucha frecuencia por nuestros participantes y, por lo tanto, es probable que dicha estructura tenga un nivel de activación relativamente bajo en las dos primeras lenguas de los aprendientes. Así, podemos considerar que los aprendientes relacionan la estructura N+N con una estructura más frecuente en ambas lenguas, la estructura N+A (que aparece en los sintagmas nominales de ambas lenguas). De esta manera, no sorprende observar que los errores de direccionalidad son poco frecuentes en la L3, ya que las dos fuentes de influencia potencial presentan la misma direccionalidad en este tipo de estructuras (N+A). En cuanto a los compuestos deverbales, la estructura parece lo suficientemente distinta como para que los participantes establezcan una relación entre el francés y el español (misma direccionalidad, morfología similar), pero no entre el euskera y el español (direccionalidad inversa, morfología distinta). Efectivamente, el CM considera que la posibilidad de influencia se basa en los indicios estructurales, por lo que, por definición, una estructura demasiado distinta, como en el caso los deverbales en euskera y en español, no será una fuente de influencia privilegiada. En términos del CM, este tipo de comparación entre estructuras distintas da lugar a un proceso que se llama diferenciación. En un contexto de adquisición de L2, esto daría lugar a dos estructuras diferenciadas, una para la L1, otra para la L2. En adquisición de L3, si hay diferenciación entre la L1 y la L3, también puede entrar en juego la estructura de la L2, a pesar de que esta tenga un nivel de activación más bajo.

4.2.4 Estrategias metalingüísticas

Otra manera de abordar la interpretación de lo observado en cuanto al género gramatical consiste en considerar las estrategias (implícitas) utilizadas por nuestros participantes, no como el resultado de una influencia interlingüística directa, sino como el fruto de ciertas habilidades metalingüísticas.

Caffarra et al. (2017) observaron que los bilingües euskera-español dominantes de euskera procesaban los indicios de género gramatical (la terminación de los sustantivos) en español de

manera distinta a los dominantes de español. Los autores sugieren que dicha diferencia de procesamiento se debe a la transferencia de una rutina sintáctica del euskera hacia el español, algo que no parece producirse cuando el español es la lengua dominante. Sin embargo, no se trataría de una transferencia directa, sino de una reutilización de una rutina sintáctica (análisis de los casos) para otro fin (análisis del género gramatical). En este sentido, este tipo de influencia parece distinto de otras influencias interlingüísticas más directas que se pueden observar en nuestro estudio, como, por ejemplo, la influencia de la direccionalidad de los compuestos del francés en los compuestos producidos en español.

A su vez, la aplicación del criterio analógico representa otra estrategia, también diferente de las demás influencias observadas. En este caso, se trata de la extensión del rasgo de género de una lengua (francés) a otra lengua (español).

En cada grupo predomina una estrategia diferente para seleccionar el género gramatical: el grupo FR aplica el criterio analógico, mientras que el grupo EU utiliza la terminación de los sustantivos. En el marco de nuestro experimento, la primera estrategia resulta menos beneficiosa, ya que nuestro diseño comporta menos ítems de género congruente (50 %) de los que el francés y el español comparten en general (alrededor del 70 %, según Teschner, 1987). Por otra parte, el uso de la terminación resulta más beneficioso ya que nuestro experimento no incluye sustantivos cuya terminación canónica (-a/-o) no corresponda al género esperado (p. ej., mano, mapa). Por eso, los hablantes del grupo EU muestran un porcentaje de aciertos más altos. No obstante, contempladas de manera global, las dos estrategias son positivas.

En ese sentido, si consideramos estas dos estrategias como el resultado de una influencia facilitadora a partir del francés y del euskera, nuestros resultados serían compatibles con el CEM, el cual siempre predice una influencia facilitadora o neutra para evitar la redundancia en el proceso de adquisición. Sin embargo, este modelo no permite explicar por qué no todos los grupos muestran la misma influencia facilitadora en la misma medida.

4.3 Reflexiones para la enseñanza de ELE

La presente tesis se enfoca en la adquisición y no en la didáctica, por lo que nuestras conclusiones no son necesariamente aplicables a la enseñanza del español como L3. Como subrayó Lightbown (1985), los hallazgos del campo de la adquisición no siempre pueden aplicarse

directamente al campo de la enseñanza. Sin embargo, creemos que nuestro estudio puede proporcionar pistas de reflexión para alimentar la discusión sobre la influencia interlingüística en la enseñanza de L3.

En cuanto al efecto del criterio analógico que hemos podido observar para el grupo FR, hay que recordar que, en el caso del francés y del español, este efecto tiende a ser positivo, ya que hay congruencia de género en alrededor del 70 % de los sustantivos equivalentes entre las dos lenguas (Teschner, 1987). Para reforzar este efecto en la enseñanza del español, se debería insistir en el hecho de que, si bien el criterio analógico funciona en la mayoría de los casos, es necesario prestar especial atención a algunas palabras de uso frecuente (nariz, cama, etc.) que no tienen un género congruente con el francés. Por otro lado, el efecto de la morfofonología observado en el grupo EU merece consideración por parte de los docentes, ya que este criterio por sí solo permite asignar correctamente el género de una mayoría de sustantivos sin recurrir a la analogía con otra lengua.

Para ejemplificar el fenómeno, hemos analizado los sustantivos que aparecen en la primera unidad de *Aula Internacional 1* (Corpas, García, et al., 2013), un manual de español de nivel inicial. La Tabla 4.2 presenta los 73 sustantivos inanimados recopilados en esta unidad separados en función de la transparencia de su terminación y de su congruencia de género con el francés.¹⁹⁴

Tabla 4.2. Análisis de los sustantivos de la Unidad 1 del manual *Aula Internacional 1*

	Congruente	Incongruente	Total
Transparente	36 (49,31 %)	8 (10,96 %)	44 (60,27 %)
	(ej. “casa”)	(ej. “ropa”)	
Opaco	28 (38,36 %)	1 (1,37 %)	29 (39,73 %)
	(ej. “dirección”)	(ej. “edad”)	
Total	64 (87,67 %)	9 (12,33 %)	73 (100 %)

¹⁹⁴ Hemos analizado solo los sustantivos inanimados para ser coherente con el diseño de nuestro estudio que no incluía sustantivos animados.

Según los datos proporcionados en la Tabla 4.2, si los aprendientes francófonos aplican el criterio analógico a los sustantivos del español, deberían acertar en el 87,76 % de los casos y si combinan las dos estrategias que recomendamos aquí (la morfofonología y el criterio analógico con el francés), solo uno de los 73 sustantivos de esta unidad debería presentar un problema. Se trata del sustantivo “edad” que es femenino en español y masculino en francés (“*âge*”) y tiene una terminación opaca. Sabiendo esto, el docente puede insistir en el género de este sustantivo potencialmente problemático y proponer indicios de concordancia con un determinante (p. ej. “la/una edad”) o un adjetivo (p. ej., “edad avanzada”, “edad media”) que conlleva una marca clara de género y evitar los determinantes y los adjetivos que carecen de tales marcas (p. ej., “su edad escolar”).

Por otra parte, como sugiere Lyster (2016) para la enseñanza del francés como L2, el docente debería incitar a sus alumnos a producir siempre los sustantivos acompañados de un determinante (p. ej., “la edad”). En cuanto a la transparencia de la terminación, los manuales, como en el caso de *Aula Internacional 1*, suelen incluir algunas explicaciones sobre las terminaciones -o/-a para los sustantivos animados (nacionalidades y profesiones), pero no amplían estas explicaciones al resto de los sustantivos que carecen de género biológico. Esta explicación podría resultar provechosa, como muestra la Tabla 4.2, ya que el 60,27 % de los sustantivos inanimados de la unidad 1 (donde aparecen las explicaciones sobre el género) tienen una terminación transparente. Si bien la muestra que analizamos es reducida (no asumimos que estos mismos porcentajes se puedan extender a todo el vocabulario enseñado en los niveles iniciales), consideramos que la reflexión merece ser profundizada en investigaciones futuras sobre la enseñanza del género gramatical en español, especialmente cuando el francés es una de las lenguas previamente adquiridas.

Con respecto a los compuestos, no queremos insistir en la enseñanza de los N+N ya que, por una parte, no parecen presentar problemas ni de direccionalidad ni de forma para nuestros participantes y, por otra, según la literatura sobre los compuestos, se trata de una estructura poco productiva en español. Sin embargo, nos parece importante hacer hincapié en los compuestos deverbales, que sí son productivos en español, porque nuestros participantes muestran problemas con la forma de estos compuestos (en la flexión verbal y en la -s genérica). A pesar de representar una estructura productiva, los compuestos deverbales no suelen formar parte de los programas de enseñanza (Liceras et al., 2004) y pasan a menudo desapercibidos en los manuales de ELE (Serrano

Dolader, 2018). A modo de ejemplo, el único compuesto deverbal que hemos encontrado en los dos primeros niveles de *Aula internacional* (Corpas, Garmendia, et al., 2013) aparece en el nivel A2 y es “lavavajillas” (p. 42). En *ELE Actual*, el equivalente “lavaplatos” aparece en el nivel A1, pero, de nuevo, se trata del único compuesto deverbal encontrado en todo el manual (Borobio, 2011, p. 85). En ambos casos, no se hace ninguna aclaración acerca de la morfología de estos compuestos. Con la poca atención que se dedica a los compuestos deverbales en los currículos y su casi ausencia en el *input* en clase, Licerias et al. (2004) afirman que es muy probable que muchos aprendientes no hayan tenido contacto alguno con estas estructuras. Consideramos que, con un foco en la forma, se podrían adquirir de manera relativamente rápida, ya que las características morfológicas de los deverbales forman parte de lo que se adquiere en los niveles iniciales del español. Concretamente, pensando en el papel del *input* en el proceso de adquisición, al dirigir la atención en la forma a la hora de trabajar los compuestos deverbales, se podría proporcionar un *input* adicional que no se relaciona solo con los deverbales, sino también con fenómenos gramaticales más generales de la lengua, como la flexión verbal, el orden de palabras y el uso del plural con valor genérico. Además, como demuestran los trabajos de Álvarez et al. (1997, 1999) sobre el español y de Nicoladis sobre el francés (2003, 2005), el uso de los deverbales para acuñar palabras nuevas representa una estrategia que emplean de manera muy creativa los niños en su LN, por lo que se podría explotar más en el proceso de adquisición de L2N, ya sea L2 o L3.

4.4 Principales aportaciones

La presente tesis permitió cumplir con los dos objetivos principales planteados en la introducción y, de estos dos objetivos, se desprenden varias aportaciones.

Nuestro primer objetivo consistía en analizar y profundizar en los aspectos metodológicos de los estudios empíricos de AL3, con especial interés en las dificultades y desafíos metodológicos propios al estudio de la adquisición multilingüe. Para ello, diseñamos un protocolo experimental que nos permitió evaluar dos fenómenos distintos en las tres lenguas de nuestros participantes. Por una parte, el análisis de dos fenómenos nos permitió comparar si la influencia se ejercía de manera global o si se podían observar diferencias dependiendo del fenómeno estudiado. Por otra parte, la administración del experimento en las tres lenguas de los hablantes nos permitió tomar en cuenta las posibles divergencias entre la descripción teórica de las lenguas y las lenguas-I de nuestros participantes. Con ello, pudimos observar que la direccionalidad de ciertos compuestos N+N en

euskera (núcleo final) no era tan clara como se suele mencionar en la literatura, lo cual nos permitió matizar la posible influencia de la direccionalidad de los compuestos del euskera en español L3. Por otra parte, descubrimos que nuestros participantes utilizaban un sufijo poco documentado (-*tzeko*) para formar los compuestos deverbales con un referente instrumental, lo que no implicó ningún cambio en nuestras hipótesis iniciales sobre la influencia hacia la L3, pero que resulta de interés para futuros estudios sobre estas estructuras en euskera.¹⁹⁵ Además, nuestros datos en francés nos permitieron resaltar la diferencia entre los compuestos deverbales con referente instrumental (poco productivos) y aquellos con un referente animado (más productivos). Si bien existían datos de corpus y trabajos descriptivos sobre esta diferencia, los datos experimentales eran escasos. Otro desafío metodológico de nuestro estudio consistía en evaluar el nivel de nuestros participantes en todas sus lenguas para poder describir su perfil lingüístico de la manera más precisa posible. Si bien esto representa un reto en sí mismo, hacerlo con una población infantil es todavía más complejo. Para cumplir con ello, adaptamos una prueba de repetición elicitada (Bowden, 2016; Kim et al., 2016; L. Ortega et al., 2002; Tracy-Ventura et al., 2014) que podrá ser reutilizada en futuros estudios tras ciertos ajustes mínimos (Lacroix & Alba de la Fuente, 2019; en preparación).

Nuestro segundo objetivo consistía en contribuir al campo de la AL3 con datos provenientes de una combinación de lenguas poco estudiada: el trío francés-euskera-español. Nuestra discusión de la sección 4.2 ilustra cómo el hecho de explorar dos fenómenos distintos con este trío de lenguas puede enriquecer el campo de la AL3. A este respecto, el hallazgo de las diferentes estrategias utilizadas en la selección del género gramatical (la terminación frente al criterio analógico) por parte de los hablantes en función de su lengua dominante es de particular interés para alimentar la discusión sobre los modelos de AL3. Además, nuestros datos permitieron subrayar el papel de un factor menos estudiado en la producción de los compuestos: el rasgo de animacidad. Por un lado, nuestros datos indican que el valor [+animado] favorece la producción de compuestos N+N y el valor [-animado] favorece la producción de deverbales. Especialmente en el caso los N+N, se trata de un factor poco estudiado (con la excepción de Liceras & Klassen, 2019) que merecería más atención. Otra innovación de nuestro estudio en relación con el campo de AL3 tiene que ver con el hecho de que nuestros participantes eran niños aprendientes de L3, algo que no se suele estudiar

¹⁹⁵ R. Etxepare (comunicación personal) apunta que el sufijo *-tzeko* está asociado a un contexto más informal, mientras que *-gailu* se relaciona con un registro más formal.

mucho ni en el campo de AL2 ni en AL3. La mayoría de los estudios de la disciplina se suelen enfocar, o bien en la adquisición bilingüe (simultánea o temprana), o en la adquisición de una LNN después de la pubertad. Los estudios sobre la adquisición del inglés como L3 por parte de niños en el País Vasco (Cenoz, 2001, 2013; Cenoz & Valencia, 1994; Valencia & Cenoz, 1992; entre otros) y en Cataluña (Portolés & Safont, 2018; entre otros) representan algunas excepciones a esta tendencia en el campo de AL3. Nuestro estudio siguió la estela de estos trabajos al interesarse por la adquisición de una LNN por parte de niños antes de la pubertad.

De manera más general, nuestro trabajo también presenta aportaciones más allá del campo de la adquisición. Consideramos que la descripción gramatical detallada de los dos fenómenos estudiados en las tres lenguas representa una aportación en sí misma y que el acercamiento teórico mixto que adoptamos nos permitió observar dichos fenómenos desde una multiplicidad de perspectivas y, al mismo tiempo, fomentar el diálogo entre los acercamientos basados en el uso y el generativismo.

4.5 Investigaciones futuras

Si bien la combinación de lenguas estudiadas en la presente tesis permitió arrojar luz sobre la influencia de las lenguas previamente adquiridas en el género gramatical en español L3, otras combinaciones de lenguas podrían complementar nuestros hallazgos. Por ejemplo, sería interesante aplicar el mismo protocolo con participantes que hablen, en lugar de euskera, otra lengua sin género y sin una morfología tan rica (p. ej., el inglés). Eso nos permitiría discernir mejor entre los conocimientos metalingüísticos procedentes de la adquisición de la L2 y el impacto de tener una lengua aglutinante como L1. Este tipo de investigación se podría llevar a cabo en la provincia de Quebec o en otras regiones bilingües de Canadá (p. ej. Ottawa) donde los aprendientes de español suelen tener el francés y el inglés como sus dos primeras lenguas. Por otra parte, para observar de manera más directa la atención que prestan los aprendientes a la terminación de los sustantivos, una posibilidad de metodología interesante sería el uso del seguimiento ocular (*eye-tracking*) con un experimento adaptado.

Para seguir indagando en el papel del criterio analógico, sería de interés también aplicar nuestro protocolo a un trío de lenguas que presente las mismas categorías de género (proximidad estructural), pero que presente divergencias con respecto a la tipología. Un ejemplo de esta

combinación sería el trío francés, hebreo y español, tres lenguas que comportan los géneros masculino y femenino, pero que pertenecen a familias lingüísticas diferentes (romance en el caso del español y el francés y semítica en el caso del hebreo). El hecho de observar un efecto de la lengua tipológicamente más lejana permitiría desambiguar entre la proximidad estructural y la tipológica.

Pasando ahora a los compuestos, particularmente a los N+N, la aplicación de nuestro protocolo a niños que hablan francés como L1 y que adquieren el español como L2 nos proporcionaría una perspectiva de interés para comparar nuestros datos. Al averiguar si estos niños presentan una tasa de productividad de esta estructura comparable a nuestros participantes de español L3, podríamos comprobar si el hecho de tener el euskera como lengua previamente adquirida fomenta la producción de compuestos en la L3. Si bien Puig-Mayenco et al. (2020) argumentan que los grupos espejo (como en nuestro estudio, L1 francés/L2 euskera frente a L1 euskera/L2 francés) representan el diseño óptimo para los estudios sobre AL3, consideramos que comparar datos de L3 con datos de L2 también puede ofrecer también información valiosa para el campo de AL3.

Finalmente, considerando la escasez de estudios sobre la adquisición de los compuestos en euskera, podría ser de interés replicar estudios llevados a cabo con niños bilingües francés-inglés (Krott et al., 2009, 2010; Murphy & Nicoladis, 2006; Nicoladis, 1999, 2002a, 2002b, 2002c, 2006) con grupos de niños bilingües francés-euskera y español-euskera. Esto nos permitiría observar, por una parte, si los compuestos en euskera se adquieren de manera comparable a lo observado para los hablantes de inglés y, por otra, comprobar si hay variación en el uso de los compuestos en euskera en función de su situación de contacto con el francés o con el español. También para llegar a un mayor nivel de precisión de descripción de los compuestos en euskera, sería interesante aplicar nuestro protocolo a grupos de control de adultos bilingües euskera-francés dominantes en cada una de las dos lenguas para comprobar los fenómenos menos documentados en la literatura que hemos observado en nuestros datos, como el uso del sufijo *-tzeke* para formar compuestos deverbales instrumentales y la direccionalidad variable en los compuestos N+N.

Bibliografía

- Abney, S. P. (1987). *The English noun phrase in its sentential aspect* [Tesis doctoral]. Massachusetts Institute of Technology.
- Acuña-Fariña, J. C., Meseguer, E., & Carreiras, M. (2014). Gender and number agreement in comprehension in Spanish. *Lingua*, *143*, 108-128.
<https://doi.org/10.1016/j.lingua.2014.01.013>
- Adger, D. (2003). *Core syntax: A minimalist approach*. Oxford University Press.
- Alarcón, A., Irma Verónica. (2011). Spanish gender agreement under complete and incomplete acquisition: Early and late bilinguals' linguistic behavior within the noun phrase. *Bilingualism: Language and Cognition*, *14*(03), 332-350.
<https://doi.org/10.1017/S1366728910000222>
- Alarcón, I. V. (2005). *The second language acquisition of Spanish gender agreement: The effects of linguistic variables on accuracy* [Tesis doctoral, Indiana University].
<https://search.proquest.com/docview/304986413/abstract/3AC0F226A01A4C9APQ/1>
- Alba de la Fuente, A., Cruz Enríquez, M., & Lacroix, H. (2018). Mood Selection in Relative Clauses by French–Spanish Bilinguals: Contrasts and Similarities between L2 and Heritage Speakers. *Languages*, *3*(3), 31. <https://doi.org/10.3390/languages3030031>
- Alba de la Fuente, A., & Lacroix, H. (2015). Multilingual learners and foreign language acquisition: Insights into the effects of prior linguistic knowledge. *Language Learning in Higher Education*, *5*(1), 45-57. <https://doi.org/10.1515/cercles-2015-0003>
- Alberdi, J. (1995). The development of the Basque system of terms of address and the allocutive conjugation. En J. I. Hualde, J. Lakarra, & R. L. Trask (Eds.), *Towards a history of the Basque language* (pp. 275-293). John Benjamins Publishing Company.

- Alcaraz Varó, E., & Martínez Linares, M. A. (1997). *Diccionario de lingüística moderna*. Editorial Ariel.
- Álvarez, A., Casares, M. F., Olivares, M. A., & Zinkgraf, M. (1997). Compuestos deverbales instrumentales del español. *Revista de Lengua y literatura*, 9-11(17-22), 117-128.
- Álvarez, A., Casares, M. F., Olivares, M. A., & Zinkgraf, M. (1999). La adquisición de compuestos en el español: ¿un proceso tardío? *Revista de Logopedia, Foniatría y Audiología*, 19(4), 190-202. [https://doi.org/10.1016/S0214-4603\(99\)75728-1](https://doi.org/10.1016/S0214-4603(99)75728-1)
- Álvarez Benito, G., & Iñigo Mora, I. M. (2005). Estrategias discursivo-kinésicas de la evasión. *AnMal electrónica*, 18, 1-16.
- Andersen, R. (1984). What is Gender Good For, Anyway? En R. Andersen (Ed.), *Second languages: A cross-linguistic perspective* (pp. 77-99). Newbury House Publishers.
- Arnaud, P. J. L. (2003). *Les composés timbre-poste*. Presses universitaires de Lyon.
- Arnaud, P. J. L. (2016). Categorizing the modification relations in French relational subordinative [NN]Ncompounds. En P. ten Hacken (Ed.), *The Semantics of Compounding* (pp. 71-93). Cambridge University Press.
<https://doi.org/10.1017/CBO9781316163122.005>
- Arnaud, P. J. L. (2018). Bateau phare, magasin phare: Composés [N1N2]N et séquences syntaxiques N1+N2 à N2 adjectivé. *Travaux de linguistique*, n° 76(1), 7-26.
- Arnaud, P. J. L., & Renner, V. (2014). English and French [NN]N lexical units: A categorial, morphological and semantic comparison. *Word Structure*, 7(1), 1-28.
- Artiagoitia, X. (1991). Aspects of tenseless relative clauses in Basque. *Anuario del Seminario de Filología Vasca «Julio de Urquijo» (ASJU)*, 25(3), 697-712.
- Artiagoitia, X. (2002). -(T)ar atzizkidun hitzen jokabide sintaktikoaren inguruan. *Fontes Linguae Vasconum*, 91, 443-462.

- Artiagoitia, X., Hualde, J. I., & Urbina, J. O. de. (2016). Basque. En P. O. Müller, I. Ohnheiser, S. Olsen, & F. Rainer (Eds.), *Word-Formation*. De Gruyter.
<https://doi.org/10.1515/9783110424942-014>
- Ayoun, D. (2010). Corpus data: Shedding the light on French grammatical gender ... or not. *EUROSLA Yearbook*, 10(1), 119-141. <https://doi.org/10.1075/eurosla.10.08ayo>
- Ayoun, D. (2018). Grammatical gender assignment in French: Dispelling the native speaker myth. *Journal of French Language Studies*, 28(1), 113-148.
<https://doi.org/10.1017/S095926951700014X>
- Azkarate, M. (1993). Basque Compound Nouns and Generative Morphology: Some data. En J. I. Hualde & J. Ortiz de Urbina (Eds.), *Generative studies in Basque linguistics* (pp. 221-242). John Benjamins Publishing Company.
- Azkarate, M. (1995). Sobre los compuestos de tipo sustantivo + sust/adj deverbal en euskera. *ASJUren Gehigarriak*, 38, 39-50.
- Azkarate Villar, M. (1990). *Hitz elkartuak euskaraz*. Deustoko Unibertsitatea.
- Azpiroz, M. I., & Ezeizabarrena Seguro, M. J. (2012). La resolución anafórica en la adquisición del euskera. *Empiricism and analytical tools for 21 Century applied linguistics: selected papers from the XXIX International Conference of the Spanish Association of Applied Linguistics (AESLA)*, 185, 65.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390-412.
<https://doi.org/10.1016/j.jml.2007.12.005>
- Badiola, L., & Sande, A. (2018). Gender assignment in Basque/Spanish mixed determiner phrases: A study of simultaneous bilinguals. En L. López (Ed.), *Issues in Hispanic and*

Lusophone Linguistics (Vol. 19, pp. 15-38). John Benjamins Publishing Company.

<https://doi.org/10.1075/ihl1.19.02bad>

Baker, M. C. (2008). The macroparameter in a microparametric world. En T. Biberauer (Ed.), *Linguistik Aktuell/Linguistics Today* (Vol. 132, pp. 351-373). John Benjamins Publishing Company. <https://doi.org/10.1075/la.132.16bak>

Bange, P. (2002). L'usage de la règle dans l'enseignement et l'apprentissage de langue étrangère. En F. Cicurel & D. Veronique (Eds.), *Discours, action et appropriation des langues* (pp. 21-36). Presses Sorbonne Nouvelle.

Barbaud, P. (1997). Composition lexicale et nominalisation: Où est la morphologie ? *Sillexicales*, 1, 25-34.

Barbaud, P., Ducharme, Ch., & Valois, D. (1982). D'un usage particulier du genre en canadien-français: La féminisation des noms à initiale vocalique. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, 27(02), 103-133.

<https://doi.org/10.1017/S0008413100023860>

Barber, H., & Carreiras, M. (2005). Grammatical Gender and Number Agreement in Spanish: An ERP Comparison. *Journal of Cognitive Neuroscience*, 17(1), 137-153.

<https://doi.org/10.1162/0898929052880101>

Barbosa, P. G., & Nicoladis, E. (2016). Deverbal compound comprehension in preschool children. *The Mental Lexicon*, 11(1), 94-114. <https://doi.org/10.1075/ml.11.1.05bar>

Bardel, C., & Falk, Y. (2007). The role of the second language in third language acquisition: The case of Germanic syntax. *Second Language Research*, 23(4), 459-484.

<https://doi.org/10.1177/0267658307080557>

Bardel, C., & Sánchez, L. (2017). The L2 status factor hypothesis revisited: The role of metalinguistic knowledge, working memory, attention and noticing in third language

- learning. En T. Angelovska & A. Hahn (Eds.), *Bilingual Processing and Acquisition* (Vol. 5, pp. 85-101). John Benjamins Publishing Company. <https://doi.org/10.1075/bpa.5.05bar>
- Barreña, A. (1997). Desarrollo diferenciado de sistemas gramaticales en un niño vasco-español bilingüe. En A. T. Pérez-Leroux & W. R. Glass (Eds.), *Contemporary perspectives on the acquisition of Spanish* (Vol. 1, pp. 55-74).
- Bartoń, K. (2019). *MuMIn: Multi-model inference*. <https://CRAN.R-project.org/package=MuMIn>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48.
<https://doi.org/10.18637/jss.v067.i01>
- Bates, E., & MacWhinney, B. (1987). Competition, variation, and language learning. En *Mechanisms of language acquisition* (pp. 157-193). Lawrence Erlbaum Associates Publishers.
- Bauer, L. (1978). *The Grammar of Nominal Compounding with Special Reference to Danish, English, and French*. Odense University Press.
- Bauer, L. (2001). Compounding. En M. Haspelmath (Ed.), *Language Typology and Language Universals* (Vol. 1, pp. 695-707). Walter de Gruyter.
- Bayram, F., Kupisch, T., Pascual y Cabo, D., & Rothman, J. (2019). Terminology matters on theoretical grounds too!: Coherent grammars cannot be incomplete. *Studies in Second Language Acquisition*, 41(2), 257-264. <https://doi.org/10.1017/S0272263119000287>
- Beard, R. (1995). *Lexeme-morpheme base morphology: A general theory of inflection and word formation*. State University of New York Press.
- Beard, R. (1996). *Base Rule Ordering of Heads in Nominal Compounds*.
- Bedore, L. M., Peña, E. D., Summers, C. L., Boerger, K. M., Resendiz, M. D., Greene, K., Bohman, T. M., & Gillam, R. B. (2012). The measure matters: Language dominance

- profiles across measures in Spanish–English bilingual children. *Bilingualism: Language and Cognition*, 15(3), 616-629. <https://doi.org/10.1017/S1366728912000090>
- Benczes, R. (2006). *Creative compounding in English: The semantics of metaphorical and metonymical noun-noun combinations*. John Benjamins Publishing Company.
- Benveniste, E. (1974). Formes nouvelles de la composition nominale. En *Problèmes de linguistique générale* (Vol. 2, pp. 163-176). Gallimard. (Original work published 1966)
- Bergen, J. J. (1978). A Simplified Approach for Teaching the Gender of Spanish Nouns. *Hispania*, 61(4), 865-876. <https://doi.org/10.2307/340934>
- Berkes, É., & Flynn, S. (2012). Further evidence in support of the Cumulative-Enhancement Model: CP structure development. En J. Cabrelli Amaro, S. Flynn, & J. Rothman (Eds.), *Third Language Acquisition in Adulthood* (pp. 143-164). John Benjamins Publishing Company. <https://doi.org/10.1075/sibil.46.11ber>
- Berman, R. A. (2009). Children's acquisition of compound constructions. En R. Lieber & P. Štekauer (Eds.), *The Oxford handbook of compounding* (pp. 298-322). Oxford University Press.
- Berman, R. A., & Clark, E. V. (1989). Learning to use compounds for contrast: Data from Hebrew. *First Language*, 9(27), 247-270.
- Berro Urrizelki, A., Fernández, B., & Ortiz de Urbina, J. (Eds.). (2019). *Basque and Romance: Aligning grammars*. Brill.
- Bialystok, E., & Barac, R. (2013). Cognitive Effects. En F. Grosjean & P. Li, *Psycholinguistics of Bilingualism* (pp. 192-213).
- Bialystok, E., Craik, F. I. M., & Luk, G. (2012). Bilingualism: Consequences for mind and brain. *Trends in Cognitive Sciences*, 16(4), 240-250. <https://doi.org/10.1016/j.tics.2012.03.001>

- Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*, 13(4), 525-531. <https://doi.org/10.1017/S1366728909990423>
- Birdsong, D. (2006). Age and Second Language Acquisition and Processing: A Selective Overview. *Language Learning*, 56(s1), 9-49. <https://doi.org/10.1111/j.1467-9922.2006.00353.x>
- Bisetto, A., & Scalise, S. (2005). The Classification of Compounds. *Lingue e Linguaggio*, IV(2), 319-332.
- Bley-Vroman, R. (1989). What is the logical problem of foreign language learning? En S. M. Gass & J. Schachter (Eds.), *Linguistic Perspectives on Second Language Acquisition* (pp. 41-68). Cambridge University Press. <https://doi.org/10.1017/CBO9781139524544.005>
- Bloomfield, L. (1973). *Language*. Compton Printing. (Original work published 1933)
- Boloh, Y., & Ibernou, L. (2013). Natural gender, phonological cues and the default grammatical gender in French children. *First Language*, 33(5), 449-468. <https://doi.org/10.1177/0142723713499848>
- Bonaparte, L.-L. (1862). *Langue basque et langues finnoises*. Strangeways & Walden. <https://gallica.bnf.fr/ark:/12148/bpt6k8851970>
- Bongaerts, T. (1999). Ultimate attainment in L2 pronunciation: The case of very advanced late L2 learners. En D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 133-159). Lawrence Erlbaum Associates Publishers. <https://www.taylorfrancis.com/books/e/9781410601667/chapters/10.4324/9781410601667-11>
- Booij, G. (2005). Compounding and Derivation: Evidence for Construction Morphology. En W. U. Dressler, D. Kastovsky, O. E. Pfeiffer, & F. Rainer (Eds.), *Current Issues in Linguistic*

- Theory* (Vol. 264, pp. 109-132). John Benjamins Publishing Company.
<https://doi.org/10.1075/cilt.264.08boo>
- Booij, G. (2009). Compounding and Construction Morphology. En R. Lieber & P. Štekauer (Eds.), *The Oxford handbook of compounding* (pp. 322-430). Oxford University Press.
- Booij, G. (2010). Compound construction: Schemas or analogy?: A construction morphology perspective. En S. Scalise & I. Vogel (Eds.), *Current Issues in Linguistic Theory* (Vol. 311, pp. 93-108). John Benjamins Publishing Company.
<https://doi.org/10.1075/cilt.311.09boo>
- Booij, G. (2013). Morphology in Construction Grammar. En T. Hoffmann & G. Trousdale (Eds.), *The Oxford Handbook of Construction Grammar* (Vol. 1). Oxford University Press.
<https://doi.org/10.1093/oxfordhb/9780195396683.013.0014>
- Borobio, V. (2011). *ELE Actual A1: Curso de español para extranjeros. Libro del alumno*. S.M. Internacional.
- Boroditsky, L., Schmidt, L. A., & Phillips, W. (2003). Sex, Syntax, and Semantics. En D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 61-79). MIT Press.
- Bosque, I., & Gutiérrez-Rexach, J. (Eds.). (2009). *Fundamentos de sintaxis formal*. Akal.
- Bottari, P., Cipriani, P., & Chilosi, A. M. (1993). Protosyntactic Devices in the Acquisition of Italian Free Morphology. *Language Acquisition*, 3(4), 327-369.
https://doi.org/10.1207/s15327817la0304_1
- Bowden, H. W. (2016). Assessing second-language oral proficiency for research: The Spanish Elicited Imitation Task. *Studies in Second Language Acquisition*, 38(4), 647-675.
<https://doi.org/10.1017/S0272263115000443>

- Bowden, H. W., Sanz, C., & Stafford, C. A. (2005). Individual differences: Age, sex, working memory, and prior knowledge. En C. Sanz (Ed.), *Mind and context in adult second language acquisition: Methods, theory, and practice* (pp. 105-140).
<http://www.jstor.org/stable/j.ctt2tt6xc.8>
- Boxell, O. (2016). The Place of Universal Grammar in the Study of Language and Mind: A Response to Dabrowska (2015). *Open Linguistics*, 2(1), 352-372.
<https://doi.org/10.1515/opli-2016-0017>
- Brinton, D. M., Kagan, O., & Bauckus, S. (2008). *Heritage Language Education: A New Field Emerging*. Routledge.
- Bruhn de Garavito, J., & White, L. (2002). The second language acquisition of Spanish DPs: The status of grammatical features. En A. T. Pérez-Leroux & J. M. Liceras (Eds.), *The Acquisition of Spanish Morphosyntax* (Vol. 31). Springer. <https://doi.org/10.1007/978-94-010-0291-2>
- Buenaftuentes de La Mata, C. (2017). Morfología léxica, variación lingüística y lexicografía: La composición léxica nominal en el *Diccionario de americanismos*. *Bulletin of Hispanic Studies*, 94(7), 677-696. <https://doi.org/10.3828/bhs.2017.42>
- Bull, W. E. (1965). *Spanish for teachers: Applied linguistics*. Ronald Press Co.
<https://archive.org/details/spanishforteach00bull>
- Burstall, C., Cohen, S., & Hargreaves, M. (1974). *Primary French in the balance*. NFER Publishing Company.
- Butt, J., & Benjamin, C. (2011). *A new reference grammar of modern Spanish* (5th ed). Oxford University Press.

- Bybee, J. (2008). Usage-based Grammar and Second Language Acquisition. En P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 216-236). Routledge.
- Bybee, J. (2010). *Language, Usage and Cognition*. Cambridge University Press.
<https://doi.org/10.1017/CBO9780511750526>
- Cabrelli Amaro, J. (2012). L3 phonology: An understudied domain. En J. Cabrelli Amaro, S. Flynn, & J. Rothman (Eds.), *Third language acquisition in adulthood* (pp. 33-60). Benjamins. <https://doi.org/10.1075/sibil.46.05ama>
- Cabrelli Amaro, J., Amaro, J. F., & Rothman, J. (2015). The relationship between L3 transfer and structural similarity across development. En H. Peukert (Ed.), *Transfer effects in multilingual language development* (Vol. 4, pp. 21-52).
<https://doi.org/10.1075/hsl.4.02ama>
- Caffarra, S., & Barber, H. A. (2015). Does the ending matter? The role of gender-to-ending consistency in sentence reading. *Brain Research*, 1605, 83-92.
<https://doi.org/10.1016/j.brainres.2015.02.018>
- Caffarra, S., Barber, H., Molinaro, N., & Carreiras, M. (2017). When the end matters: Influence of gender cues during agreement computation in bilinguals. *Language, Cognition and Neuroscience*, 32(9), 1069-1085. <https://doi.org/10.1080/23273798.2017.1283426>
- Carroll, S. (1989). Second-Language Acquisition and the Computational Paradigm. *Language Learning*, 39(4), 535-594. <https://doi.org/10.1111/j.1467-1770.1989.tb00902.x>
- Carroll, S. E., & Widjaja, E. (Nina). (2013). Learning exponents of number on first exposure to an L2. *Second Language Research*, 29(2), 201-229.
<https://doi.org/10.1177/0267658312473471>
- Carstens, V. (2000). Concord in Minimalist Theory. *Linguistic Inquiry*, 31(2), 319-355.

- Carstens, V. (2001). Multiple Agreement and Case Deletion: Against phi-incompleteness. *Syntax*, 4(3), 147-163. <https://doi.org/10.1111/1467-9612.00042>
- Carstens, V. (2010). Implications of grammatical gender for the theory of uninterpretable features. En M. T. Putnam (Ed.), *Exploring crash-proof grammars*. John Benjamins Publishing Company.
- Carvalho, A. M., & Silva, A. J. B. (2006). Cross-Linguistic Influence in Third Language Acquisition: The Case of Spanish-English Bilinguals' Acquisition of Portuguese. *Foreign Language Annals*, 39(2), 185-202. <https://doi.org/10.1111/j.1944-9720.2006.tb02261.x>
- Cenoz, J. (2001). The effect of linguistic distance, L2 status and age on cross-linguistic influence in L3 acquisition. En J. Cenoz, B. Hufeisen, & U. Jessner (Eds.), *Cross-linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives* (pp. 8-20). Multilingual Matters.
- Cenoz, J. (2003a). The additive effect of bilingualism on third language acquisition: A review. *International Journal of Bilingualism*, 7(1), 71-87. <https://doi.org/10.1177/13670069030070010501>
- Cenoz, J. (2003b). The influence of age on the acquisition of English: General proficiency, attitudes and code mixing. En M. del P. García Mayo & M. L. García Lecumberri (Eds.), *Age and the acquisition of English as a foreign language* (pp. 77-93). Multilingual Matters.
- Cenoz, J. (2013). The influence of bilingualism on third language acquisition: Focus on multilingualism. *Language Teaching*, 46(1), 71-86. <https://doi.org/10.1017/S0261444811000218>

- Cenoz, J., Hufeisen, B., & Jessner, U. (2001). *Cross-linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives*. Multilingual Matters.
<https://doi.org/10.21832/9781853595509>
- Cenoz, J., & Valencia, J. F. (1994). Additive trilingualism: Evidence from the Basque Country. *Applied Psycholinguistics*, 15(2), 195-207. <https://doi.org/10.1017/S0142716400005324>
- Chomsky, N. (1959). Verbal behavior [Review of *Verbal behavior*, por B. F. Skinner]. *Language*, 35(1), 26-58. JSTOR. <https://doi.org/10.2307/411334>
- Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.
- Chomsky, N. (1980). *Rules and representations*. Columbia University Press.
- Chomsky, N. (1981). *Lectures on government and binding*. Foris.
- Chomsky, N. (1982). *Some concepts and consequences of the theory of government and binding*. MIT Press.
- Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. Praeger.
- Chomsky, N. (1995). *The minimalist program*. MIT Press.
- Chomsky, N. (1999). Derivation by phase. *MIT occasional papers in linguistics*, 18.
- Chomsky, N. (2000). Minimalist inquiries: The framework. En R. Martin, D. Michaels, & J. Uriagareka (Eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik* (pp. 89-155). MIT Press.
- Chomsky, N. (2001). Derivation by phase. En J. Kenstowicz (Ed.), *Ken Hale: A life in language* (pp. 1-52). MIT Press.
- Chomsky, N. (2002). *Syntactic structures*. Mouton. (Original work published 1957)
- Chomsky, N. (2007). Approaching UG from below. En U. Sauerland & H.-M. Gärtner (Eds.), *Interfaces + recursion = language? Chomsky's minimalism and the view from syntax-semantics* (pp. 1-29). Mouton de Gruyter.

- Chomsky, N. (2008). On phases. En R. Freidin, C. P. Otero, & M. L. Zubizarreta (Eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud* (pp. 133-136). MIT Press.
- Christiansen, M. H., & Chater, N. (2015). The language faculty that wasn't: A usage-based account of natural language recursion. *Frontiers in Psychology*, 6, 1182.
<https://doi.org/10.3389/fpsyg.2015.01182>
- Christiansen, M. H., & MacDonald, M. (2009). A Usage-Based Approach to Recursion in Sentence Processing. *Language as a Complex Adaptive System*, 59(s1), 126-161.
<https://doi.org/10.1111/j.1467-9922.2009.00538.x>
- Cinque, G. (1993). A Null Theory of Phrase and Compound Stress. *Linguistic Inquiry*, 24(2), 239-297. JSTOR.
- Clahsen, H., & Felser, C. (2006a). Continuity and shallow structures in language processing. *Applied Psycholinguistics*, 27(1), 107-126. <https://doi.org/10.1017/S0142716406060206>
- Clahsen, H., & Felser, C. (2006b). Grammatical processing in language learners. *Applied Psycholinguistics*, 27(1), 3-42. <https://doi.org/10.1017/S0142716406060024>
- Clark, E. V. (1985). The acquisition of Romance with special reference to French. En D. I. Slobin (Ed.), *The crosslinguistic of language acquisition* (Vol. 1, pp. 687-782). Lawrence Erlbaum Associates Publishers.
- Clark, E. V. (1998). Lexical creativity in French-speaking children. *Cahiers de psychologie cognitive*, 17(2), 513-530.
- Clark, E. V. (2009). *First language acquisition* (2.^a ed.). Cambridge University Press.
- Clark, E. V., & Berman, R. A. (1987). Types of linguistic knowledge: Interpreting and producing compound nouns. *Journal of Child language*, 14(3), 547-567.

- Clark, E. V., Frant Hecht, B., & Mulford, R. C. (1986). Coining complex compounds in English: Affixes and word order in acquisition. *Linguistics*, 24(1), 7-29.
<https://doi.org/10.1515/ling.1986.24.1.7>
- Clark, E. V., Gelman, S. A., & Lane, N. M. (1985). Compound Nouns and Category Structure in Young Children. *Child Development*, 56(1), 84-94.
- Collentine, J., & Freed, B. F. (2004). Learning context and its effects on second language acquisition: Introduction. *Studies in Second Language Acquisition*, 26(2), 153-171.
<https://doi.org/10.1017/S0272263104262015>
- Comrie, B. (1989). *Language Universals and Linguistic Typology* (2.^a ed.). University of Chicago Press.
- Comrie, B. (1999). Grammatical Gender Systems: A Linguist's Assessment. *Journal of Psycholinguistic Research*, 28(5), 457-466. <https://doi.org/10.1023/A:1023212225540>
- Comrie, B., Haspelmath, M., & Bickel, B. (2004). *Leipzig Glossing Rules*.
<https://www.eva.mpg.de/lingua/resources/glossing-rules.php>
- Conseil supérieur de la langue française. (1990). Les rectifications de l'orthographe. *Journal officiel de la République française. Document administratif.*, 100.
- Consejo de Europa. (2002). *Marco común de referencia para las lenguas: Aprendizaje, enseñanza y evaluación*. (Instituto Cervantes, Trad.). Secretaría General Técnica del MECD-Subdirección General de Información y Publicaciones, y Grupo Anaya.
http://cvc.cervantes.es/ensenanza/biblioteca_ele/marco/cvc_mer.pdf
- Cook, V. (1991). The poverty-of-the-stimulus argument and multicompetence. *Second Language Research*, 7(2), 103-117. <https://doi.org/10.1177/026765839100700203>
- Cook, V. (1992). Evidence for Multicompetence. *Language Learning*, 42(4), 557-591.
<https://doi.org/10.1111/j.1467-1770.1992.tb01044.x>

- Cook, V., & Li, W. (2016). *The Cambridge handbook of linguistic multi-competence*.
<https://doi.org/10.1017/CBO9781107425965>
- Corbett, G. G. (1991). *Gender*. Cambridge University Press.
<https://doi.org/10.1017/CBO9781139166119>
- Corbett, G. G. (2006a). *Agreement*. Cambridge University Press.
- Corbett, G. G. (2006b). Gender, Grammatical. En K. Brown (Ed.), *Encyclopedia of Language & Linguistics (Second Edition)* (pp. 749-756). Elsevier. <https://doi.org/10.1016/B0-08-044854-2/00191-7>
- Corbett, G. G., Fraser, N. M., & McGlashan, S. (Eds.). (1993). *Heads in grammatical theory*. Cambridge University Press.
- Corbin, D. (1992). Hypothèses sur les frontières de la composition nominale. *Cahiers de grammaire*, 17, 25-55.
- Corpas, J., García, E., & Garmendia, A. (2013). *Aula internacional 1. Nueva edición. Libro del alumno A1* (2.^a ed.). Difusión.
- Corpas, J., Garmendia, A., & Soriano, C. (2013). *Aula internacional 2. Nueva edición. Libro del alumno A2* (Nueva edición). Difusión.
- Council of Europe. (2018). *Common European Framework of Reference for Languages: Learning, teaching, assessment companion volume with new descriptors*.
<https://rm.coe.int/cefr-companion-volume-with-new-descriptors-2018/1680787989>
- Coyos, J.-B. (2004). Basque. En P. J. L. Arnaud (Ed.), *Le nom composé: Données en seize langues* (pp. 47-70). Presses universitaires de Lyon.
- Croft, W. (1991). *Syntactic categories and grammatical relations: The cognitive organization of information*. University of Chicago Press.

- Croft, W. (2001). *Radical construction grammar: Syntactic theory in typological perspective*. Oxford University Press.
- Cruse, D. A. (1986). *Lexical semantics*. Cambridge Univ. Press.
- Cuza, A., & Pérez-Tattam, R. (2016). Grammatical gender selection and phrasal word order in child heritage Spanish: A feature re-assembly approach. *Bilingualism: Language and Cognition*, 19(01), 50-68.
- Da Silva, Z. S. (1975). *A concept approach to Spanish*. Harper and Row.
- Dąbrowska, E. (2015). What exactly is Universal Grammar, and has anyone seen it? *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.00852>
- Darmesteter, A. (1894). *Traité de la formation des mots composés dans la langue française comparée aux autres langues romanes et au latin*. 2. Éd., vue, corr., et en partie refondue, avec une préf. Par Gaston Paris. Paris E. Bouillon.
<http://archive.org/details/traitdelaforma00darmuoft>
- Davies, M. (2019). *Corpus del Español NOW corpus (News on the Web)*.
<https://www.corpusdelespanol.org/now/>
- De Angelis, G. (2005). Multilingualism and Non-native Lexical Transfer: An Identification Problem. *International Journal of Multilingualism*, 2(1), 1-25.
<https://doi.org/10.1080/17501220508668374>
- De Angelis, G. (2007). *Third or additional language acquisition*. Multilingual Matters.
<https://doi.org/10.21832/9781847690050>
- de Bot, K. (2004). The Multilingual Lexicon: Modelling Selection and Control. *International Journal of Multilingualism*, 1(1), 17-32. <https://doi.org/10.1080/14790710408668176>
- de Bustos Gisbert, E. (1986). *La composición nominal en español*. Universidad de Salamanca.

- De Cat, C., Klepousniotou, E., & Baayen, R. H. (2015). Representational deficit or processing effect? An electrophysiological study of noun-noun compound processing by very advanced L2 speakers of English. *Frontiers in Psychology*, 6.
<https://doi.org/10.3389/fpsyg.2015.00077>
- De Houwer, A. (1990). *The acquisition of two languages from birth: A case study*. Cambridge University Press.
- De Houwer, A. (2007). Parental language input patterns and children's bilingual use. *Applied Psycholinguistics*, 28(3). <https://doi.org/10.1017/S0142716407070221>
- De Houwer, A., Bornstein, M. H., & Putnick, D. L. (2014). A bilingual–monolingual comparison of young children's vocabulary size: Evidence from comprehension and production. *Applied Psycholinguistics*, 35(6), 1189-1211.
<https://doi.org/10.1017/S0142716412000744>
- de Rijk, R. P. G. (2008). *Standard Basque: A progressive grammar*. MIT Press.
- Desmets, M., & Villoing, F. (2010). Morphologie constructionnelle et arguments sémantiques du verbe: Un traitement HPSG des composés VN du français. *Travaux de linguistique*, 60(1), 65. <https://doi.org/10.3917/tl.060.0065>
- Desrochers, A., Licerias, J. M., Spradlin, T., & Fuertes, R. F. (2003). Can an On-line Response Latency Task Shed Light on Native and Non-native Competence in the Deverbal Compounds of Spanish? En J. M. Licerias, H. Zobl, & H. Goodluck (Eds.), *Proceedings of the 6th Generative Approaches to Second Language Acquisition Conference (GASLA 2002): L2 Links* (p. 9). Cascadilla Proceedings Project.
- Dewaele, J.-M. (2015). Gender Errors in French Interlanguage: The Effect of Initial Consonant Versus Initial Vowel of the Head Noun. *Arborescences : Revue d'études Françaises*, 5, 7-27. <https://doi.org/10.7202/1032661ar>

- Di Sciullo, A. M. (1991). On the structure of deverbal compounds. *Working Papers in Linguistics, 1*, 1-26.
- Di Sciullo, A. M. (1992). Deverbal compounds and the external argument. En I. M. Roca (Ed.), *Thematic Structure*. De Gruyter Mouton. <https://doi.org/10.1515/9783110872613.65>
- Di Sciullo, A. M., & Williams, E. (1987). *On the definition of word*. MIT Press.
- Dijkstra, T. (2003). Lexical processing in bilinguals and multilinguals: The word selection problem. En J. Cenoz, B. Hufeisen, & U. Jessner (Eds.), *The Multilingual Lexicon* (pp. 11-26). Kluwer Academic Publishers.
- Dinnes, I. S. (1971). Must All Unclassified Spanish Words Be Memorized for Gender? *Hispania, 54*(3), 487-492. <https://doi.org/10.2307/338378>
- Domínguez, L., Hicks, G., & Slabakova, R. (2019). Terminology choice in generative acquisition research: The case of “incomplete acquisition” in heritage language grammars. *Studies in Second Language Acquisition, 41*(2), 241-255.
<https://doi.org/10.1017/S0272263119000160>
- Dowens, M. G., Guo, T., Guo, J., Barber, H., & Carreiras, M. (2011). Gender and number processing in Chinese learners of Spanish – Evidence from Event Related Potentials. *Neuropsychologia, 49*(7), 1651-1659.
<https://doi.org/10.1016/j.neuropsychologia.2011.02.034>
- Drackert, A. (2015). *Validating language proficiency assessments in second language acquisition research: Applying an argument-based approach*. Peter Lang D.
<https://doi.org/10.3726/978-3-653-06280-9>
- Dressler, W. U. (2006). Compound types. En G. Libben & G. Jarema (Eds.), *The representation and processing of compound words* (pp. 23-44). Oxford University Press.

- Dryer, M. S. (2013). Order of adjective and noun. En M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online*. Max Planck Institute for Evolutionary Anthropology. <https://wals.info/chapter/87>
- Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: One-stop shopping for Spanish word properties. *Behavior research methods*, 45(4), 1246-1258. <https://doi.org/10.3758/s13428-013-0326-1>
- Duffield, N., & White, L. (1999). Assessing L2 knowledge of Spanish clitic placement: Converging methodologies. *Second Language Research*, 15(2), 133-160. <https://doi.org/10.1191/026765899668237583>
- Duffield, N., White, L., Garavito, J. B. D., Montrul, S., & Prévost, P. (2002). Clitic Placement in L2 French: Evidence from Sentence Matching. *Journal of Linguistics*, 38(3), 487-525. JSTOR.
- Duguine, I. (2013). *Le développement du langage oral chez l'enfant bilingue basque-français en contexte d'acquisition simultanée vs. successive des langues: Étude de la maîtrise du cas ergatif en basque et du genre grammatical en français* [Tesis doctoral, Toulouse 2]. <http://www.theses.fr/2013TOU20008>
- Duguine, I., Köpke, B., & Nespoulous, J.-L. (2014). Variations dans l'acquisition du marqueur ergatif en basque par des enfants bilingues basque-français. *Language, Interaction and Acquisition*, 5(2), 227-251.
- Echeverría, B. (2011). Indexing religious identity in the French Basque Country: Toward a theory of pronominal shift. *Anuario Del Seminario de Filología Vasca «Julio de Urquijo»*, 40(1-2), 273-291.
- Eguren, L. (1989). Algunos datos del español en favor de la hipótesis de la frase determinante. *Revista Argentina de Lingüística*, 5.1/2, 163-209.

- Eguren, L. (2009). El morfema de alocutivo del euskera y el modelo de gramática. *Hermēneus. Revista de Traducción e Interpretación*, 0(2), 95-118.
- Elhuyar Fundazioa. (2018). *Elhuyar Hiztegiak*. Elhuyar Fundazioa - Hizkuntza eta Teknologia Unitatea. <https://hiztegiak.elhuyar.eus/>
- Ellis, N. C. (1999). Cognitive approaches to SLA. *Annual Review of Applied Linguistics*, 19, 22-42. <https://doi.org/10.1017/S0267190599190020>
- Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24(2), 143-188. <https://doi.org/10.1017/S0272263102002024>
- Ellis, N. C. (2006). Selective Attention and Transfer Phenomena in L2 Acquisition: Contingency, Cue Competition, Saliency, Interference, Overshadowing, Blocking, and Perceptual Learning. *Applied Linguistics*, 27(2), 164-194. <https://doi.org/10.1093/applin/aml015>
- Ellis, N. C., & Cadierno, T. (2009). Constructing a Second Language: Introduction to the Special Section. *Annual Review of Cognitive Linguistics*, 7(1), 111-139. <https://doi.org/10.1075/arcl.7.05ell>
- Ellis, N. C., & Collins, L. (2009). Input and Second Language Acquisition: The Roles of Frequency, Form, and Function Introduction to the Special Issue. *The Modern Language Journal*, 93(3), 329-335. <https://doi.org/10.1111/j.1540-4781.2009.00893.x>
- Epstein, S. D., Flynn, S., & Martohardjono, G. (1996). Second language acquisition: Theoretical and experimental issues in contemporary research. *Behavioral and Brain Sciences*, 19(4), 677-714.
- Eubank, L. (1993). On the Transfer of Parametric Values in L2 Development. *Language Acquisition*, 3(3), 183-208. https://doi.org/10.1207/s15327817la0303_1

- Eubank, L. (1996). Negation in early German-English Interlanguage: More Valueless Features in the L2 initial state. *Second Language Research*, 12(1), 73-106.
<https://doi.org/10.1177/026765839601200104>
- European Commission. (2007). *Final report: High Level Group on Multilingualism*. Office for Official Publications of the European Communities. <https://op.europa.eu/s/oe1F>
- Euskaltzaindia/Real Academia de la Lengua Vasca. (1992). *Hitz-Elkarketa/4. Hitz elkartuen osaera eta idazkera [Composición y ortografía de las palabras compuestas]*. Euskaltzaindia.
- Eusko Jaurlaritzza - Gobierno Vasco. (2019). *VI Encuesta Sociolingüística*. Eusko Jaurlaritzaren Argitalpen Zerbitzu Nagusia / Servicio Central de Publicaciones del Gobierno Vasco.
https://www.habe.euskadi.eus/contenidos/noticia/eas_mas_noticias/es_def/adjuntos/INKE_STA_ESt.pdf
- Evans, V. (2014). *The language myth: Why language is not an instinct*. Cambridge University Press.
- Everett, D. L. (2016). An evaluation of Universal Grammar and the phonological mind. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.00015>
- Ezeizabarrena, M. J. (2013). Overt subjects in early Basque and other null subject languages. *International Journal of Bilingualism*, 17(3), 309-336.
<https://doi.org/10.1177/1367006912438997>
- Ezeizabarrena, M.-J. (2012). The (in) consistent ergative marking in early Basque: L1 vs. child L2. *Lingua*, 122(3), 303-317.
- Ezeizabarrena Segurola, M. J. (1996). *Adquisición de la morfología verbal en euskera y español por niños bilingües*. Servicios de Publicaciones de la Universidad del País Vasco.

- Ezeizabarrena Segurola, M. J., Garate, I. M., & Lizarralde, L. B. (2009). Euskara H2 goiztiarraren ezaugarrien bila: Adizkiak eta gramatika-kasuak haurren ipuin-kontaketetan. *Euskera: Euskaltzaindiaren lan eta agiriak*= *Trabajos y actas de la Real Academia de la Lengua Vasca*= *Travaux et actes de l'Academie de la Langue basque*, 54(2), 639-681.
- Fabb, N. (2001). Compounding. En A. Spencer & A. M. Zwicky (Eds.), *The handbook of morphology* (4.^a ed., pp. 66-83). John Wiley & Sons.
- Fábregas, A., & Masini, F. (2015). Prominence in morphology: The notion of head. *Lingue e linguaggio*, XIV(1), 79-96.
- Falk, Y., & Bardel, C. (2010). The study of the role of the background languages in third language acquisition. The state of the art. *IRAL - International Review of Applied Linguistics in Language Teaching*, 48(2-3), 185-219.
<https://doi.org/10.1515/iral.2010.009>
- Falk, Y., & Bardel, C. (2011). Object pronouns in German L3 syntax: Evidence for the L2 status factor. *Second Language Research*, 27(1), 59-82.
<https://doi.org/10.1177/0267658310386647>
- Falk, Y., Lindqvist, C., & Bardel, C. (2015). The role of L1 explicit metalinguistic knowledge in L3 oral production at the initial state. *Bilingualism: Language and cognition*, 18(2), 227-235.
- Faussart, C., Jakubowicz, C., & Costes, M. (1999). Gender and number processing in spoken French and Spanish. *Rivista Di Linguistica*, 75-101.
- Fernández, B., Oihartzabal, B., Landa, J., & Sarasola, I. (2010). -Le/-tzaile inflexiozko atzizkiaz. *Anuario del Seminario de Filología Vasca «Julio de Urquijo» (ASJU)*, 44(2), 457-510.
- Fernández Fuertes, R., Álvarez de la Fuente, E., & Mujcinovic, S. (2016). The acquisition of grammatical gender in L1 bilingual Spanish. En A. Alba de la Fuente, E. Valenzuela, &

C. Martínez Sanz (Eds.), *Language Acquisition Beyond Parameters: Studies in honour of Juana M. Liceras* (pp. 237-279). John Benjamins Publishing Company.

<https://doi.org/10.1075/sibil.51.10fue>

Fernández-García, M. (1999). Patterns of gender agreement in the speech of second language learners. En J. Gutiérrez-Rexach & F. Martínez-Gil (Eds.), *Advances in Hispanic linguistics: Papers from the 2nd Hispanic linguistics symposium*. Cascadilla Press.

Fillmore, C. J. (1988). The Mechanisms of “Construction Grammar”. *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*, 35-55.

<https://doi.org/10.3765/bls.v14i0.1794>

Finnemann, M. D. (1992). Learning Agreement in the Noun Phrase: The Strategies of Three First-Year Spanish Students. *IRAL - International Review of Applied Linguistics in Language Teaching*, 30(2), 121-136. <https://doi.org/10.1515/iral.1992.30.2.121>

Flege, J. E. (1987). *Effects of equivalence classification on the production of foreign language speech sounds* (J. Allan & J. Leather, Eds.). Foris Publications.

Flynn, S. (1983). *A Study of the Effects of Principal Branching Direction in Second Language Acquisition: The Generalization of a Parameter of Universal Grammar from First to Second Language Acquisition* [Tesis doctoral]. Cornell University.

Flynn, S. (1987). *A parameter-setting model of L2 acquisition: Experimental studies in anaphora*. Reidel.

Flynn, S., Foley, C., & Vinnitskaya, I. (2004). The Cumulative-Enhancement Model for Language Acquisition: Comparing Adults' and Children's Patterns of Development in First, Second and Third Language Acquisition of Relative Clauses. *International Journal of Multilingualism*, 1(1), 3-16. <https://doi.org/10.1080/14790710408668175>

- Flynn, S., & Lust, B. (1980). Acquisition of relative clauses in English: Developmental changes in their head. En *Cornell Working Papers in Linguistics* (Vol. 1, pp. 33-45). Department of Modern Languages and Linguistics, Cornell University.
- Foote, R. (2014). Age of Acquisition and Sensitivity to Gender in Spanish Word Recognition. *Language Acquisition*, 21(4), 365-385. <https://doi.org/10.1080/10489223.2014.892948>
- Fradin, B. (2005). On a semantically grounded difference between derivation and compounding. En W. U. Dressler, D. Kastovsky, O. E. Pfeifer, & F. Rainer (Eds.), *Morphology and its demarcations: Selected papers from the 11th Morphology Meeting, Vienna, February 2004* (pp. 161-182). John Benjamins Pub. Co.
- Fradin, B. (2009). IE, Romance: French. En R. Lieber & P. Štekauer (Eds.), *The Oxford handbook of compounding* (pp. 659-689). Oxford University Press.
- Franceschina, F. (2001). Morphological or syntactic deficits in near-native speakers? An assessment of some current proposals. *Second Language Research*, 17(3), 213-247. <https://doi.org/10.1177/026765830101700301>
- Franceschina, F. (2005). *Fossilized Second Language Grammars: The acquisition of grammatical gender*. John Benjamins Publishing Company. <https://doi.org/10.1075/lald.38>
- García Lecumberri, M. L., & Gallardo, F. (2003). English FL sounds in school learners of different ages. En M. del P. García Mayo & M. L. García Lecumberri (Eds.), *Age and the acquisition of English as a foreign language* (pp. 115-135). Multilingual Matters.
- García Mayo, M. del P. (2003). Age, length of exposure and grammaticality judgements in the acquisition of English as a foreign. En M. del P. García Mayo & M. L. García Lecumberri (Eds.), *Age and the acquisition of English as a foreign language* (pp. 94-114). Multilingual Matters.

- García Mayo, M. del P. (2006). Synthetic Compounding in the English Interlanguage of Basque–Spanish Bilinguals. *International Journal of Multilingualism*, 3(4), 231-257.
- García Mayo, M. del P. (2012). Cognitive approaches to L3 acquisition. *International Journal of English Studies*, 12(1), 129-146. <https://doi.org/10.6018/ijes.12.1.140421>
- García Mayo, M. del P., Ibarrola, A. L., & Licerias, J. M. (2006). Agreement in the English Interlanguage of Basque/Spanish Bilinguals: A minimalist farewell to *pro*. *ITL - International Journal of Applied Linguistics*, 151(1), 83-98.
<https://doi.org/10.2143/ITL.151.0.2015223>
- García Mayo, M. del P., & Olaizola, I. V. (2011). The development of suppletive and affixal tense and agreement morphemes in the L3 English of Basque-Spanish bilinguals. *Second Language Research*, 27(1), 129-149. <https://doi.org/10.1177/0267658310386523>
- García Mayo, M. del P., & Slabakova, R. (2015). Object drop in L3 acquisition. *International Journal of Bilingualism*, 19(5), 483-498. <https://doi.org/10.1177/1367006914524643>
- García-Page, M. (2011). Hombre clave, hombre rana, ¿un mismo fenómeno? *Verba. Anuario Galego de Filoloxía*, 38, 44.
- Garza-González, P. (2014). *The Interpretation of N+N and V+N Compounds by Spanish Heritage Speakers* [Tesis doctoral, Texas A&M University].
<https://oaktrust.library.tamu.edu/handle/1969.1/153877>
- Gass, S. M. (1988). Second Language Acquisition and Linguistic Theory: The Role of Language Transfer. En S. Flynn & W. O'Neil (Eds.), *Linguistic Theory in Second Language Acquisition* (pp. 384-403). Springer Netherlands. https://doi.org/10.1007/978-94-009-2733-9_21

- Gathercole, V. C. M., Sebastián, E., & Soto, P. (1999). The early acquisition of Spanish verbal morphology: Across-the-board or piecemeal knowledge? *International Journal of Bilingualism*, 3(2-3), 133-182. <https://doi.org/10.1177/13670069990030020401>
- Gillon Dowens, M., Vergara, M., Barber, H. A., & Carreiras, M. (2009). Morphosyntactic Processing in Late Second-Language Learners. *Journal of Cognitive Neuroscience*, 22(8), 1870-1887. <https://doi.org/10.1162/jocn.2009.21304>
- Giurescu, A. (1975). *Les mots composés dans les langues romanes*. Mouton.
- Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. University of Chicago Press.
- Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford University Press.
- Gollan, T. H., & Frost, R. (2001). Two Routes to Grammatical Gender: Evidence from Hebrew. *Journal of Psycholinguistic Research*, 30(6), 627-651.
- Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always a means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of memory and language*, 58(3), 787-814. <https://doi.org/10.1016/j.jml.2007.07.001>
- González Alonso, J., & Rothman, J. (2016). Coming of age in L3 initial stages transfer models: Deriving developmental predictions and looking towards the future. *International Journal of Bilingualism*, 21(6). <https://doi.org/10.1177/1367006916649265>
- González Alonso, J., Villegas, J., & García Mayo, M. del P. (2016). English compound and non-compound processing in bilingual and multilingual speakers: Effects of dominance and sequential multilingualism. *Second Language Research*, 32(4), 503-535. <https://doi.org/10.1177/0267658316642819>

- Gordon, P. (1985). Level-ordering in lexical development. *Cognition*, 21(2), 73-93.
[https://doi.org/10.1016/0010-0277\(85\)90046-0](https://doi.org/10.1016/0010-0277(85)90046-0)
- Granfeldt, J. (2005). The Development of Gender Attribution and Gender Agreement in French: A Comparison of Bilingual First and Second Language Learners. En J.-M. Dewaele (Ed.), *Focus on French as a Foreign Language: Multidisciplinary Approaches* (pp. 164-190). Multilingual Matters. <https://doi.org/10.21832/9781853597688>
- Gregg, K. R. (2003). The state of emergentism in second language acquisition. *Second Language Research*, 19(2), 95-128. <https://doi.org/10.1191/0267658303sr213oa>
- Grégoire, A. (1947). *L'Apprentissage du langage*. Duculot.
- Grevisse, M., & Goosse, A. (2008). *Le bon usage: Grammaire française ; Grevisse langue française* (14.^a ed.). De Boeck-Duculot.
- Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, 36(1), 3-15.
- Grosjean, F. (2010). *Bilingual*. Harvard University Press.
- Grosjean, F., & Li, P. (2013). *The psycholinguistics of bilingualism*. Wiley-Blackwell.
- Grüter, T., Lew-Williams, C., & Fernald, A. (2012). Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research*, 28(2), 191-215.
<https://doi.org/10.1177/0267658312437990>
- Güemes, M., Gattei, C., & Wainelboim, A. (2019). Processing verb–noun compound words in Spanish: Evidence from event-related potentials. *Cognitive Neuropsychology*, 36(5-6), 265-281. <https://doi.org/10.1080/02643294.2019.1618254>
- Guevara, E., & Scalise, S. (2009). Searching for Universals in Compounding. En S. Scalise, E. Magni, & A. Bisetto (Eds.), *Universals of language today* (1. ed). Springer.

- Gullberg, M., Roberts, L., Dimroth, C., Veroude, K., & Indefrey, P. (2010). Adult Language Learning After Minimal Exposure to an Unknown Natural Language. *Language Learning*, 60(s2), 5-24. <https://doi.org/10.1111/j.1467-9922.2010.00598.x>
- Gutiérrez Mangado, M. J., & Martínez-Adrián, M. (2018). CLIL at the linguistic interfaces. *Journal of Immersion and Content-Based Language Education*, 6(1), 85-112. <https://doi.org/10.1075/jicb.17002.gut>
- Gutiérrez Rodriguez, V. (2015). *Reconocimiento y generación de palabras compuestas en español* [Tesis doctoral, Universidad de Las Palmas de Gran Canaria]. <http://hdl.handle.net/10553/22927>
- Gutiérrez-Mangado, M. J., & Martínez-Adrián, M. (2018). The use of L3 English articles by Basque-Spanish bilinguals. *Revista Española de Lingüística Aplicada/Spanish Journal of Applied Linguistics*, 31(1), 124-158. <https://doi.org/10.1075/resla.17015.gut>
- Hammarberg, B. (2010). The languages of the multilingual: Some conceptual and terminological issues. *IRAL - International Review of Applied Linguistics in Language Teaching*, 48(2-3), 91-104. <https://doi.org/10.1515/iral.2010.005>
- Hammarberg, B. (2018). L3, the tertiary language. En A. Bonnet & P. Siemund (Eds.), *Hamburg Studies on Linguistic Diversity* (Vol. 7, pp. 127-150). John Benjamins Publishing Company. <https://doi.org/10.1075/hslld.7.06ham>
- Han, Z. (2004). *Fossilization in Adult Second Language Acquisition*. Multilingual Matters.
- Han, Z. (2013). Forty years later: Updating the Fossilization Hypothesis. *Language Teaching*, 46(2), 133-171. <https://doi.org/10.1017/S0261444812000511>
- Han, Z., & Tarone, E. (Eds.). (2014). *Interlanguage: Forty years later* (Vol. 39). John Benjamins Publishing Company. <https://doi.org/10.1075/lllt.39>

- Harbour, D., Adger, D., & Béjar, S. (Eds.). (2008). *Phi theory: Phi-features across modules and interfaces*. Oxford University Press.
- Harris, J. W. (1987). The accentual patterns of verb paradigms in Spanish. *Natural Language and Linguistic Theory*, 5(1), 61-90. <https://doi.org/10.1007/BF00161868>
- Harris, J. W. (1991). The Exponence of Gender in Spanish. *Linguistic Inquiry*, 22(1), 27-62.
- Harris, T., & Wexler, K. (1996). The optional-infinitive stage in child English. Evidence from negation. En H. Clahsen (Ed.), *Generative Perspectives in Language Acquisition* (pp. 1-42). John Benjamins Publishing Company.
- Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science*, 298(5598), 1569-1579. <https://doi.org/10.1126/science.298.5598.1569>
- Hawkins, R., & Chan, C. Y. (1997). The partial availability of Universal Grammar in second language acquisition: The 'failed functional features hypothesis'. *Second Language Research*, 13(3), 187-226.
- Haznedar, B., & Schwartz, B. D. (1997). Are there optional infinitives in child L2 acquisition. *Proceedings of the 21st annual Boston University conference on language development*, 21, 257-268.
- Hermas, A. (2010). Language acquisition as computational resetting: Verb movement in L3 initial state. *International Journal of Multilingualism*, 7(4), 343-362. <https://doi.org/10.1080/14790718.2010.487941>
- Hermas, A. (2014). Multilingual transfer: L1 morphosyntax in L3 English. *International Journal of Language Studies*, 8(2), 1-24.

- Hernandez, A. E., Bates, E., & Avila, L. X. (1994). On-line sentence interpretation in Spanish–English bilinguals: What does it mean to be “in between”? *Applied Psycholinguistics*, 15(4), 417-446. <https://doi.org/10.1017/S014271640000686X>
- Hernández Pina, F. (1984). *Teorías psicosociolingüísticas y su aplicación a la adquisición del español como lengua materna*. Siglo XXI.
- Heyvaert, L. (2009). Compounding in Cognitive Linguistics. En R. Lieber & P. Štekauer (Eds.), *The Oxford handbook of compounding* (pp. 374-430). Oxford University Press.
- Hilpert, M., & Östman, J.-O. (2014). Reflections on Constructions across Grammars. *Constructions and Frames*, 6(2), 137-142. <https://doi.org/10.1075/cf.6.2.01int>
- Hjelmslev, L. (1959). Animé et inanimé, personnel et non-personnel. En *Essais linguistiques: Vol. XII* (pp. 211-249). Nordisk Sprog-og Kulturforlag.
- Hockett, C. F. (1958). *A course in modern linguistics*. Oxford & IBH Publishing.
- Hoekstra, T., & Hyams, N. (1998). Aspects of root infinitives. *Lingua*, 106(1), 81-112. [https://doi.org/10.1016/S0024-3841\(98\)00030-8](https://doi.org/10.1016/S0024-3841(98)00030-8)
- Hoffmann, C. (2014). *Introduction to Bilingualism*. Routledge.
- Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous Inference in General Parametric Models. *Biometrical Journal*, 50(3), 346-363.
- Hualde, J. I. (2003). Case and number inflection of noun phrases. En J. I. Hualde & J. Ortiz de Urbina (Eds.), *A grammar of Basque* (pp. 171-186). Mouton de Gruyter.
- Hualde, J. I. (2007). Stress removal and stress addition in Spanish. *Journal of Portuguese Linguistics*, 6, 59-89. <https://doi.org/10.5334/jpl.145>
- Hualde, J. I., & Ortiz de Urbina, J. (Eds.). (2003). *A grammar of Basque*. Mouton de Gruyter.
- Hufeisen, B. (1998). L3-Stand der Forschung—Was bleibt zu tun? [L3-State of the art—What needs to be done?]. En B. Hufeisen & B. Lindemann (Eds.), *L2-L3 und ihre*

zweisprachliche Interaktion: Zu individueller mehrsprachigkeit und gesteuertem lernen [L2-L3 and their crosslinguistic interaction: About individual multilingualism and instructed learning (pp. 169-183). Stauffenburg.

Hulk, A., & Tellier, C. (1999). Conflictual agreement in romance nominals. En J.-M. Authier, B. E. Bullock, & L. A. Reed (Eds.), *Formal perspectives of Romance linguistics: Selected papers from the 28th Linguistic Symposium on Romance Languages (LSRL XXVIII): University Park, 16-19 April 1998* (pp. 179-195). John Benjamins Publishing Company.

Hulstijn, J. H. (2005). Theoretical and Empirical Issues in the Study of Implicit and Explicit Second-Language Learning. *Studies in Second Language Acquisition*, 27(2), 129-140.
<https://doi.org/10.1017/S0272263105050084>

Hulstijn, J. H. (2012). The construct of language proficiency in the study of bilingualism from a cognitive perspective. *Bilingualism: Language and Cognition*, 15(2), 422-433.
<https://doi.org/10.1017/S1366728911000678>

Hulstijn, J. H., & De Graaff, R. (1994). Under what conditions does explicit knowledge of a second language facilitate the acquisition of implicit knowledge? A research proposal. *Aila Review*, 11, 97-112. <https://doi.org/11245/1.428191>

Hyltenstam, K., & Abrahamsson, N. (2003). Maturation Constraints in SLA. En C. Doughty & M. H. Long (Eds.), *The Handbook of Second Language Acquisition* (pp. 538-588). John Wiley & Sons. <https://doi.org/10.1002/9780470756492.ch17>

Idiázabal, I. (1995). First stages in the acquisition of noun phrase determiners by a Basque-Spanish bilingual child. En C. Silva-Corvalán (Ed.), *Spanish in four continents: Studies in language contact and bilingualism* (pp. 258-276). Georgetown University Press.

- Imaz Agirre, A., & García Mayo, M. del P. (2013). Gender agreement in L3 English by Basque/Spanish bilinguals. *Linguistic Approaches to Bilingualism*, 3(4), 415-447. <https://doi.org/10.1075/lab.3.4.02ima>
- Imaz Agirre, A., & García Mayo, M. del P. (2014). Double object constructions in L3 English: An exploratory study of morphological and semantic constraints. *International Journal of English Studies*, 14(2), 1-20. <https://doi.org/10.6018/j.179651>
- Jackendoff, R. (2011). Compounding in the Parallel Architecture and Conceptual Semantics. En R. Lieber & P. Štekauer (Eds.), *The Oxford Handbook of Compounding* (pp. 106-128). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199695720.013.0006>
- Jackendoff, R. (2016). English noun-noun compounds in Conceptual Semantics. En P. ten Hacken (Ed.), *The Semantics of Compounding* (pp. 15-37). Cambridge University Press. <https://doi.org/10.1017/CBO9781316163122.002>
- Jaeggli, O. (1982). *Topics in Romance syntax*. Foris.
- Jake, J. L., Myers-Scotton, C., & Gross, S. (2002). Making a minimalist approach to codeswitching work: Adding the Matrix Language. *Bilingualism: Language and Cognition*, 5(1), 69-91. <https://doi.org/10.1017/S1366728902000147>
- Jakubowicz, C., & Roulet, L. (2007). Narrow syntax or interface deficit? Gender agreement in French SLI. En J. M. Liceras, H. Zobl, & H. Goodluck (Eds.), *The Role of Formal Features in Second Language Acquisition* (pp. 185-225). Routledge.
- Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. Routledge.
- Johnson, K. (2004). What is a Spanish speaker? En A. R. M. Simões, A. M. Carvalho, & L. Wiedeman (Eds.), *Português para Falantes de Espanhol – Artigos Seleccionados Escritos*

em Português e Inglês / Portuguese for Spanish Speakers – Selected Articles Written in Portuguese and English (pp. 49-65). Pontes editores. <http://hdl.handle.net/1808/10080>

Jones, E. E., & Goethals, G. R. (1972). Order effects in impression formation: Attribution context and the nature of the entity. En E. E. Jones, D. E. Kanouse, H. H. Kelly, R. E. Nisbet, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 27-46). General Learning Press.

Kail, M., & Charvillat, A. (1988). Local and topological processing in sentence comprehension by French and Spanish children*. *Journal of Child Language*, 15(3), 637-662.
<https://doi.org/10.1017/S0305000900012605>

Karmiloff-Smith, A. (1979). Micro-and macrodevelopmental changes in language acquisition and other representational systems. *Cognitive science*, 3(2), 91-117.

Kayne, R. S. (2003). Person Morphemes and Reflexives in Italian, French, and Related Languages. En C. Tortora (Ed.), *The syntax of Italian dialects* (pp. 102-136). Oxford University Press.

Keating, G. D. (2009). Sensitivity to Violations of Gender Agreement in Native and Nonnative Spanish: An Eye-Movement Investigation. *Language Learning*, 59(3), 503-535.
<https://doi.org/10.1111/j.1467-9922.2009.00516.x>

Kellerman, E. (1983). Now you see it, now you don't. En S. M. Gass & L. Selinker (Eds.), *Language transfer in language learning* (Vol. 54, pp. 112-134). Newbury House.

Kellerman, E. (1995). Crosslinguistic influence: Transfer to nowhere? *Annual review of applied linguistics*, 15, 125-150.

Kellerman, E., & Sharwood Smith, M. (Eds.). (1986). *Crosslinguistic influence in second language acquisition* (1.^a ed.). Pergamon Institute of English.

- Kibort, A. (2010). Towards a typology of grammatical features. En A. Kibort & G. G. Corbett (Eds.), *Features: Perspectives on a key notion in linguistics* (pp. 64-106). Oxford University Press.
- Kihm, A., Cinque, G., & Kayne, R. S. (2008). Noun Class, Gender, and the Lexicon-Syntax-Morphology Interfaces: A Comparative Study of Niger-Congo and Romance Languages. En *The Oxford Handbook of Comparative Syntax*. Oxford University Press.
<https://doi.org/10.1093/oxfordhb/9780195136517.013.0011>
- Kim, Y., Tracy-Ventura, N., & Jung, Y. (2016). A Measure of Proficiency or Short-Term Memory? Validation of an Elicited Imitation Test for SLA Research. *The Modern Language Journal*, 100(3), 655-673. <https://doi.org/10.1111/modl.12346>
- Klassen, R., Liceras, J. M., & Landa-Buil, M. (2014, mayo). *On the representation of gender in the mind of the bilingual: The view from the interpretation and processing of concord and agreement code-switched structures*. 24th Colloquium on Generative Grammar, Center for Social Sciences and Humanities, Madrid, Spain.
- Klassen, R., & M. Liceras, J. (2017). The representation of gender in the mind of Spanish-English bilinguals: Insights from code-switched Adjectival Predicates. *Borealis – An International Journal of Hispanic Linguistics*, 6(1), 77. <https://doi.org/10.7557/1.6.1.4100>
- Klein, W., & Perdue, C. (1997). The Basic Variety (or: Couldn't natural languages be much simpler?). *Second Language Research*, 13(4), 301-347.
<https://doi.org/10.1191/026765897666879396>
- Koike, D., & Gualda, R. (2008). The effect of explicit or implicit teaching of grammatical form in Portuguese as a third language: Noticing and transfer. *Português para falantes de espanhol. Ensino e aquisição*, 47-67.

- Kornfeld, L. M. (2009). IE, Romance: Spanish. En R. Lieber & P. Štekauer (Eds.), *The Oxford handbook of compounding* (pp. 436-452). Oxford University Press.
- Kramer, R. (2015). *The morphosyntax of gender* (1.^a ed.). Oxford University Press.
- Krashen, S. D. (1981). *Second language acquisition and second language learning*. Pergamon Press.
- Krott, A., Gagné, C. L., & Nicoladis, E. (2009). How the parts relate to the whole: Frequency effects on children's interpretations of novel compounds. *Journal of Child Language*, 36(1), 85-112. <https://doi.org/10.1017/S030500090800888X>
- Krott, A., Gagné, C. L., & Nicoladis, E. (2010). Children's preference for HAS and LOCATED relations: A word learning bias for noun–noun compounds*. *Journal of Child Language*, 37(2), 373-394. <https://doi.org/10.1017/S0305000909009593>
- Kupisch, T., Muller, N., & Cantone, K. F. (2002). Gender in Monolingual and Bilingual First Language Acquisition: Comparing Italian and French. *Lingue e Linguaggio*, 1, 107-149. <https://doi.org/10.1418/7559>
- Kupisch, T., & Rothman, J. (2018). Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers. *International Journal of Bilingualism*, 22(5), 564-582. <https://doi.org/10.1177/1367006916654355>
- Lacroix, H. (en preparación). *A semi-automated analysis of deverbal compounds in Spanish, French and Basque dictionaries*.
- Lacroix, H., & Alba de la Fuente, A. (2019, marzo). *Sobre el uso de pruebas de imitación elicitada (EIT) como instrumento de medida de la competencia lingüística de niños multilingües* [Communication]. XV Quebec–Ontario Dialogues on Acquisition and Spanish 2019 (QODAS 2019), Ottawa, Canada.

- Lacroix, H., & Alba de la Fuente, A. (en preparación). *Elicited imitation task as a way to measure multilingual children's language proficiency.*
- Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers.* University of Michigan Press.
- Lafon, R. (1947). Sur la catégorie de genre grammatical en basque. *Bulletin hispanique*, 49(3), 373-394. <https://doi.org/10.3406/hispa.1947.3104>
- Laka, I. (1996). *A brief grammar of Euskara, the Basque language.* Euskal Herriko Unibertsitatea.
- Lakoff, G. (1987). *Women, fire, and dangerous things what categories reveal about the mind.* University of Chicago Press.
- Lakoff, G. (1990). The Invariance Hypothesis: Is abstract reason based on image-schemas? *Cognitive Linguistics (includes Cognitive Linguistic Bibliography)*, 1(1), 39-74. <https://doi.org/10.1515/cogl.1990.1.1.39>
- Lambelet, A. (2012). *L'apprentissage du genre grammatical en langue étrangère: À la croisée des approches linguistiques et cognitives* [Tesis doctoral]. Université de Fribourg.
- Lang, M. F. (1991). *Spanish word formation: Productive derivational morphology in the modern lexis.* Routledge.
- Langacker, R. W. (1987). *Foundations of cognitive grammar. Volume 1.* Stanford University Press.
- Lardiere, D. (1995). L2 acquisition of English synthetic compounding is not constrained by level-ordering (and neither, probably, is L1). *Second Language Research*, 11(1), 20-56. <https://doi.org/10.1177/026765839501100102>

- Lardiere, D. (1998). Parameter-Resetting in Morphology: Evidence from Compounding. En M.-L. Beck (Ed.), *Morphology and its interfaces in second language knowledge*. John Benjamins Publishing Company.
- Lardiere, D. (2000). Mapping features to forms in second language acquisition. En *Second language acquisition and linguistic theory* (pp. 102-129).
- Lardiere, D., & Schwartz, B. D. (1997). Feature-Marking in the L2 Development of Deverbal Compounds. *Journal of Linguistics*, 33(2), 327-353.
<https://doi.org/10.1017/S0022226797006518>
- Larrañaga, P., & Guijarro-Fuentes, P. (2013). The linguistic competence of early Basque–Spanish bilingual children and a Spanish monolingual child. *International Journal of Bilingualism*, 17(5), 577-601. <https://doi.org/10.1177/1367006911435704>
- Lemhöfer, K., Dijkstra, T., & Michel, M. (2004). Three languages, one ECHO: Cognate effects in trilingual word recognition. *Language and Cognitive Processes*, 19(5), 585-611.
<https://doi.org/10.1080/01690960444000007>
- Lenneberg, E. H. (1967). *Biological foundations of language*. Wiley.
- Leung, Y.-K. I. (2005). L2 vs. L3 initial state: A comparative study of the acquisition of French DPs by Vietnamese monolinguals and Cantonese–English bilinguals. *Bilingualism: Language and Cognition*, 8(1), 39-61. <https://doi.org/10.1017/S1366728904002044>
- Lew-Williams, C., & Fernald, A. (2007). Young Children Learning Spanish Make Rapid Use of Grammatical Gender in Spoken Word Recognition. *Psychological Science*, 18(3), 193-198. <https://doi.org/10.1111/j.1467-9280.2007.01871.x>
- Liceras, J. M. (2003). Spanish L1/L2 Crossroads: Can We Get ‘There’ from ‘Here’? En A. T. Pérez-Leroux & Y. Roberge (Eds.), *Current Issues in Linguistic Theory* (Vol. 244, pp.

317-350). John Benjamins Publishing Company.

<https://benjamins.com/catalog/cilt.244.23lic>

Liceras, J. M. (2009). La morfología léxica del español y el llamado problema lógico de la adquisición del lenguaje no nativo. En M. Zorraquino, M. Antonia, D. Serrano-Dolader, & J. F. Val Álvaro (Eds.), *Morfología y español como lengua extranjera (E/ELE)* (pp. 21-66). Prensas Universitarias de Zaragoza.

Liceras, J. M. (2010). Second Language Acquisition and Syntactic Theory in the 21st Century. *Annual Review of Applied Linguistics*, 30, 248-269.

<https://doi.org/10.1017/S0267190510000097>

Liceras, J. M. (2013, abril). *Gender agreement patterns in mixed concord and agreement structures: Does 'code-switching' matter?* Workshop on Code-Switching, Bergische Universität Wuppertal, Wuppertal, Germany.

Liceras, J. M., & Alba de la Fuente, A. (2015). Typological proximity in L2 acquisition. En T. Judy & S. Perpiñán (Eds.), *The Acquisition of Spanish in Understudied Language Pairings* (Vol. 3, pp. 329-358). John Benjamins Publishing Company.

<https://doi.org/10.1075/ihll.3>

Liceras, J. M., & Díaz, L. (2000a). La teoría chomskiana y la adquisición de la gramática no nativa: A la búsqueda de desencadenantes. En C. Muñoz (Ed.), *Segundas lenguas. Adquisición en el aula* (pp. 39-80). Ariel.

Liceras, J. M., & Díaz, L. (2000b). Triggers in L2 acquisition: The case of Spanish N-N compounds. *Studia Linguistica*, 54(2), 197-211. <https://doi.org/10.1111/1467-9582.00060>

Liceras, J. M., Díaz, L., & Salomaa-Robertson, T. (2002). The compounding parameter and the word-marker hypothesis: Accounting for the Acquisition of Spanish N-N Compounding.

- En A. T. Pérez-Leroux & J. M. Licerias (Eds.), *The acquisition of Spanish morphosyntax* (pp. 209-237). Springer.
- Licerias, J. M., Fernández Fuertes, R., Bel, A., & Martínez, C. (2012, octubre). *The mental representation of Gender and Agreement Features in Child 2L1 and Child L2 grammars: Insights from code-switching*. UIC Bilingualism Forum, University of Illinois at Chicago, Chicago, USA.
- Licerias, J. M., Fernández Fuertes, R., Perales, S., Pérez-Tattam, R., & Martínez, C. (2006, septiembre). *L2A and the pidgin/creole continuum: The pidginization/nativization hypothesis revisited*. EUROSLA, Boğaziçi University, Antalya, Turkey.
- Licerias, J. M., Fernández Fuertes, R., Perales, S., Pérez-Tattam, R., & Spradlin, K. T. (2008). Gender and gender agreement in bilingual native and non-native grammars: A view from child and adult functional–lexical mixings. *Lingua*, 118(6), 827-851.
<https://doi.org/10.1016/j.lingua.2007.05.006>
- Licerias, J. M., Fuertes, R. F., & Klassen, R. (2016). Language dominance and language nativeness: The view from English-Spanish codeswitching. *Spanish-English Codeswitching in the Caribbean and the US*, 107-138.
- Licerias, J. M., & Klassen, R. (2019). Compounding and derivation: On the ‘promiscuity’ of derivational affixes. *Linguistic Approaches to Bilingualism*, 9(1), 42-72.
<https://doi.org/10.1075/lab.16026.lic>
- Licerias, J. M., Mongeon, C., Cuza, A., Senn, C., & Spradlin, K. T. (2004). La adquisición en el aula sin input formal: Los compuestos «exocéntricos» de las interlenguas del español. *redELE: Revista Electrónica de Didáctica ELE*, 0, 17-45.

- Liceras, J. M., & Valenzuela, E. (1998, septiembre). *The compounding parameter in L2 acquisition: The subset principle revisited*. [Presentación]. Generative Approaches to Second Language Acquisition, Pittsburgh, PA.
- Liceras, J. M., Zobl, H., & Goodluck, H. (2008). *The role of formal features in second language acquisition*. Routledge.
- Lieber, R. (2016). Compounding in the lexical semantic framework. En P. ten Hacken (Ed.), *The Semantics of Compounding* (pp. 38-53). Cambridge University Press.
<https://doi.org/10.1017/CBO9781316163122.003>
- Lightbown, Patsy M. (1985). Great Expectations: Second-Language Acquisition Research and Classroom Teaching. *Applied Linguistics*, 6(2), 173-189.
- Lightbown, Patsy Martin. (1977). *Consistency and variation in the acquisition of French* [Tesis doctoral]. Columbia University.
- Lin, Y., Michel, J.-B., Aiden Lieberman, E., Orwant, J., Brockman, W., & Petrov, S. (2012). Syntactic Annotations for the Google Books Ngram Corpus. *Proceedings of the ACL 2012 System Demonstrations*, 169-174. <https://www.aclweb.org/anthology/P12-3029>
- Liu, H., Bates, E., & Li, P. (1992). Sentence interpretation in bilingual speakers of English and Chinese. *Applied Psycholinguistics*, 13(4), 451-484.
<https://doi.org/10.1017/S0142716400005762>
- Llama, R., Klassen, R., Collins, Laura, & Cardoso, W. (2011, septiembre). *It is neither “un ladder” nor “un échelle”: Mixed DPs and gender agreement in the production data of L3 Spanish learners*. EUROSLA, Stockholm University, Stockholm, Sweden.
- López Alvarado, N. E. (2009). *Encuentros*. Éditions CEC.
- López Ornat, S. (1997). What Lies in between a Pre-Grammatical and a Grammatical Representation? Evidence on Nominal and Verbal Form-Function Mappings in Spanish

- from 1;7 to 2;1. En A. T. Pérez-Leroux & W. R. Glass (Eds.), *Contemporary perspectives on the acquisition of Spanish*. (Vol. 1, pp. 3-20). Cascadilla Press.
- López Ornat, S. (2003). Learning earliest grammar: Evidence of grammar variations in speech before 22 months. En S. Montrul & F. Ordóñez (Eds.), *Linguistic theory and language development in Hispanic languages: Papers from the 5th Hispanic Linguistics Symposium and the 4th Conference on the Acquisition of Spanish and Portuguese* (pp. 254-274). Cascadilla Press.
- López Ornat, S., Mariscal, S., Fernández, A., & Gallo, P. (1994). *La adquisición de la lengua española*. <https://dialnet.unirioja.es/servlet/libro?codigo=244926>
- Lust, B., Flynn, S., Blume, M., Park, S. W., Kang, C., Yang, S., & Kim, A.-Y. (2016). Assessing child bilingualism: Direct assessment of bilingual syntax amends caretaker report. *International Journal of Bilingualism*, 20(2), 153-172.
<https://doi.org/10.1177/1367006914547661>
- Lyster, R. (2006). Predictability in French gender attribution: A corpus analysis. *Journal of French Language Studies*, 16(01), 69-92. <https://doi.org/10.1017/S0959269506002304>
- MacWhinney, B. (1987). The Competition Model. En B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 249-308). Lawrence Erlbaum Associates Publishers.
- MacWhinney, B. (1997). Second Language Acquisition and the Competition Model. En A. M. B. De Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 113-142). Lawrence Erlbaum Associates Publishers.
- MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk* (3e ed.). Lawrence Erlbaum Associates Publishers.

- MacWhinney, B. (2005). A Unified Model of Language Acquisition. En J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 49-67). Oxford University Press.
- MacWhinney, B. (2008). A Unified Model. En P. Robinson & N. C. Ellis (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition* (1.^a ed., pp. 341-371). Routledge.
<https://doi.org/10.4324/9780203938560>
- MacWhinney, B. (2012). The logic of the unified model. En *Handbook of Second Language Acquisition* (pp. 211-227). Routledge. <https://doi.org/10.4324/9780203808184.ch13>
- Maez, L. F. (1983). The Acquisition of Noun and Verb Morphology in 18–24 Month Old Spanish Speaking Children. *NABE Journal*, 7(2), 53-68.
<https://doi.org/10.1080/08855072.1983.10668442>
- Marantz, A. (1997). No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon. *University of Pennsylvania Working Papers in Linguistics*, 4(2).
<https://repository.upenn.edu/pwpl/vol4/iss2/14>
- Marchand, H. (1960). *The categories and types of present-day english word-formation: A synchronic-diachronic approach*. O. Harrassowitz.
- Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within-and between-language competition. *Bilingualism: Language and Cognition*, 6(02), 97-115. <https://doi.org/10.1017/S1366728903001068>
- Mariscal, S. (2009). Early acquisition of gender agreement in the Spanish noun phrase: Starting small. *Journal of Child Language*, 36(1), 143-171.
<https://doi.org/10.1017/S0305000908008908>
- Marqueta Gracia, B. (2017). *Restrictions in the semantic interpretation of English and Spanish compounds*. 9, 35.

- Masini, F. (2009). Phrasal lexemes, compounds and phrases: A constructionist perspective. *Word Structure*, 2(2), 254-271. <https://doi.org/10.3366/E1750124509000440>
- Mayo, M. del P. G., Ibarrola, A. L., & Licerias, J. M. (2005). Placeholders in the English Interlanguage of Bilingual (Basque/Spanish) Children. *Language Learning*, 55(3), 445-489. <https://doi.org/10.1111/j.0023-8333.2005.00312.x>
- McCarthy, C. (2008). Morphological variability in the comprehension of agreement: An argument for representation over computation. *Second Language Research*, 24(4), 459-486. <https://doi.org/10.1177/0267658308095737>
- McConnel-Ginet, S. (1988). Language and gender. En F. J. Newmeyer (Ed.), *Linguistics, the Cambridge Survey: Language: The socio-cultural context* (pp. 75-99). Cambridge University Press.
- McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, 55(3), 381-401. <https://doi.org/10.1016/j.jml.2006.06.006>
- McLaughlin, B. (1984). *Second-language acquisition in childhood* (2nd ed., Vol. 1-1 : Preschool Children). Lawrence Erlbaum Associates Publishers.
- Meisel, J. M. (1989). Early differentiation of languages in bilingual children. En K. Hyltenstam & L. K. Obler (Eds.), *Bilingualism Across the Lifespan* (pp. 13-40). Cambridge University Press. <https://doi.org/10.1017/CBO9780511611780.003>
- Meisel, J. M. (1994). Code-Switching in Young Bilingual Children: The Acquisition of Grammatical Constraints. *Studies in Second Language Acquisition*, 16(4), 413-439. <https://doi.org/10.1017/S0272263100013449>

- Meisel, J. M. (2018). Early child second language acquisition: French gender in German children. *Bilingualism: Language and Cognition*, 21(4), 656-673.
<https://doi.org/10.1017/S1366728916000237>
- Mendivil-Giró, J.-L. (2018). Is Universal Grammar ready for retirement? A short review of a longstanding misinterpretation. *Journal of Linguistics*, 54(4), 859-888.
<https://doi.org/10.1017/S0022226718000166>
- Michelena, L. (1961). *Fonética histórica vasca*. la Diputación de Guipúzcoa.
- Michelena, L. (1987). *Diccionario general vasco—Orotariko euskal hiztegia*. Desclée De Brouwer.
- Milton, J. (2009). *Measuring Second Language Vocabulary Acquisition*. Multilingual Matters.
- Montrul, S. (2004). *The Acquisition of Spanish: Morphosyntactic development in monolingual and bilingual L1 acquisition and adult L2 acquisition*. John Benjamins Publishing Company. <https://doi.org/10.1075/lald.37>
- Montrul, S. (2012). Is the heritage language like a second language? *Eurosla Yearbook*, 12(1), 1-29. <https://doi.org/10.1075/eurosla.12.03mon>
- Montrul, S. (2013). La adquisición de la lengua en la infancia. En *El bilingüismo en el mundo hispanohablante* (pp. 153-170). Wiley-Blackwell.
- Montrul, S., Foote, R., & Perpiñán, S. (2008). Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning*, 58(3), 503-553. <https://doi.org/10.1111/j.1467-9922.2008.00449.x>
- Moreno, A. I. (2010). Las estrategias de comunicación oral en español de los estudiantes neerlandófonos en Bélgica: Un estudio de caso. *Ianua. Revista Philologica Romanica*, 10, 123-141.

- Morgia, F. L. (2016). Assessing the relationship between input and strength of language development: A study on Italian–English bilingual children. En C. Silva-Corvalan & J. Treffers-Daller (Eds.), *Language Dominance in Bilinguals* (pp. 195-218). Cambridge University Press. <https://doi.org/10.1017/CBO9781107375345.010>
- Moyer, M. G. (1993). *Analysis of code-switching in Gibraltar* [Tesis doctoral]. Universitat Autònoma de Barcelona.
- Moyna, M. I. (2000). *Compounding in Spanish: Patterns and changes* [Tesis doctoral, University of Florida]. <http://archive.org/details/compoundinginspa00moyn>
- Moyna, M. I. (2006). Can we make heads or tails of Spanish endocentric compounds? *Linguistics*, 42(3), 617-637. <https://doi.org/10.1515/ling.2004.020>
- Moyna, M. I. (2011). *Compound words in Spanish: Theory and history*. John Benjamins Publishing Company.
- Müller, S., & Wechsler, S. (2014). Lexical approaches to argument structure. *Theoretical Linguistics*, 40(1-2), 1-76. <https://doi.org/10.1515/tl-2014-0001>
- Murphy, V. A. (2000). Compounding and the Representation of L2 Inflectional Morphology. *Language Learning*, 50(1), 153-197. <https://doi.org/10.1111/1467-9922.00114>
- Murphy, V. A., & Nicoladis, E. (2006). When answer-phone makes a difference in children's acquisition of English compounds. *Journal of Child Language*, 33(3), 677-691. <https://doi.org/10.1017/S030500090600746X>
- Na Ranong, S., & Leung, Y. I. (2009). Null objects in L1 Thai-L2 English-L3 Chinese: An empirical take on a theoretical problem. En Y. I. Leung (Ed.), *Third Language Acquisition and Universal Grammar* (pp. 162-191). Multilingual Matters.

- Nagpal, J., & Nicoladis, E. (2009). Why are Noun-Verb-*er* compounds so difficult for English-speaking children? *The Mental Lexicon*, 4(2), 276-301.
<https://doi.org/10.1075/ml.4.2.05nag>
- Naiman, N. (1974). The use of elicited imitation in second language acquisition research. *Working Papers on Bilingualism*, 1-37.
- Namiki, T. (1994). Subheads of compounds. *Synchronic and Diachronic Approaches to Language: A Festschrift for Toshio Nakao on the Occasion of His Sixtieth Birthday*, 269-285.
- Nathan, G. S. (1987). On second-language acquisition of voiced stops. *Journal of Phonetics*, 15(4), 313-322. [https://doi.org/10.1016/S0095-4470\(19\)30583-2](https://doi.org/10.1016/S0095-4470(19)30583-2)
- Nicoladis, E. (1999). “Where is my brush-teeth?” Acquisition of compound nouns in a French–English bilingual child. *Bilingualism: Language and Cognition*, 2(3), 245-256.
<https://doi.org/10.1017/S1366728999000346>
- Nicoladis, E. (2002a). When Is a Preposition a Linking Element? Bilingual Children’s Acquisition of French Compound Nouns. *Folia Linguistica*, 36(1-2).
<https://doi.org/10.1515/flin.2002.36.1-2.45>
- Nicoladis, E. (2002b). The Cues That Children Use in Acquiring Adjectival Phrases and Compound Nouns: Evidence from Bilingual Children. *Brain and Language*, 81(1), 635-648. <https://doi.org/10.1006/brln.2001.2553>
- Nicoladis, E. (2002c). What’s the difference between ‘toilet paper’ and ‘paper toilet’? French–English bilingual children’s crosslinguistic transfer in compound nouns. *Journal of Child Language*, 29(4), 843-863. <https://doi.org/10.1017/S0305000902005366>

- Nicoladis, E. (2003). Cross-linguistic transfer in deverbal compounds of preschool bilingual children. *Bilingualism: Language and Cognition*, 6(1), 17-31.
<https://doi.org/10.1017/S1366728903001019>
- Nicoladis, E. (2005). The Acquisition of Complex Deverbal Words by a French-English Bilingual Child. *Language Learning*, 55(3), 415-443. <https://doi.org/10.1111/j.0023-8333.2005.00311.x>
- Nicoladis, E. (2006). Preschool Children's Acquisition of Compounds. En G. Libben & G. Jarema (Eds.), *The representation and processing of compound words* (pp. 96-124). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199228911.003.0005>
- Noailly, M. (1990). *Le substantif épithète* (1re éd). Presses universitaires de France.
- Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge University Press.
- Odlin, T. (2005). Crosslinguistic influence and conceptual transfer: What are the concepts? *Annual review of applied linguistics*, 25, 3-25.
<https://doi.org/10.1017/S0267190505000012>
- Odlin, T. (2008). Cross-Linguistic Influence. En *The Handbook of Second Language Acquisition* (pp. 436-486). John Wiley & Sons. <https://doi.org/10.1002/9780470756492.ch15>
- Office Québécois de Langue Française [OQLF]. (2018). *Banque de dépannage linguistique—Amour*. http://bdl.oqlf.gouv.qc.ca/bdl/gabarit_bdl.asp?id=3489
- O'Grady, W. (1997). *Syntactic Development*. University of Chicago Press.
- O'Grady, W. (2005). *How Children Learn Language*. Cambridge University Press.
- O'Grady, W. (2008a). Does Emergentism Have a Chance? En H. Chan, H. Jacob, & E. Kiparsky (Eds.), *BUCLD 32: Proceedings of the 32nd annual Boston University Conference on Language Development* (Vol. 1, pp. 16-35). Cascadia Press.

- O'Grady, W. (2008b). Innateness, universal grammar, and emergentism. *Lingua*, 118(4), 620-631. <https://doi.org/10.1016/j.lingua.2007.03.005>
- O'Grady, W. (2010). An Emergentist Approach to Syntax. En H. Narrog & B. Heine (Eds.), *The Oxford Handbook of Linguistic Analysis* (pp. 257-283). Oxford University Press.
- Orduña-López, J. L. (2019). ¿Locuciones nominales o compuestos sintagmáticos? A propósito del español del Valle de Aburrá. *Íkala, Revista de Lenguaje y Cultura*, 24(3), 620-636. <https://doi.org/10.17533/udea.ikala.v24n03a02>
- Ortega, L., Iwashita, N., Norris, J. M., & Rabie, S. (2002, octubre). An investigation of elicited imitation tasks in crosslinguistic SLA research. *Conference Handout from Paper Presented at the Meeting of the Second Language Research Forum*.
- Ortega, M. (2008). Cross-linguistic influence in multilingual language acquisition: The role of L1 and non-native languages in English and Catalan oral production. *Íkala*, 13(19), 121-142.
- Ortiz de Urbina, J. (2019). Word Order. En A. Berro Urrizelki, B. Fernández, & J. Ortiz de Urbina (Eds.), *Basque and Romance: Aligning grammars* (pp. 14-58). Brill. https://doi.org/10.1163/9789004395398_003
- Osterhout, L., Mclaughlin, J., Kim, A., Greenwald, R., & Inoue, K. (2004). Sentences in the Brain: Event-Related Potentials as Real-Time Reflections of Sentence Comprehension and Language Learning. En M. Carreiras & Jr. C. Clifton (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERP, and beyond* (pp. 271-308).
- Otheguy, R. (2016). The linguistic competence of second-generation bilinguals: A critique of “incomplete acquisition”. En C. Tortora, M. den Dikken, I. L. Montoya, & T. O'Neill (Eds.), *Romance Languages and Linguistic Theory* (Vol. 9, pp. 301-319). John Benjamins Publishing Company. <https://doi.org/10.1075/rllt.9.16oth>

- Otheguy, R., & Lapidus Shin, N. (2003). An adaptive approach to noun gender in New York contact Spanish. En R. Núñez-Cedeño, L. López, & R. Cameron (Eds.), *A Romance perspective on language, knowledge and use: Selected papers from the 31st Linguistic Symposium on Romance Languages (LSRL), Chicago, 19-22 April 2001* (Vol. 238, pp. 209-229). John Benjamins Publishing Company. <https://doi.org/10.1075/cilt.238.17oth>
- Oyharçabal, B. (1993). Verb Agreement with Non Arguments: On allocutive agreement. En J. I. Hualde & J. Ortiz de Urbina (Eds.), *Generative studies in Basque linguistics*. John Benjamins Publishing Company.
- Oyharçabal, B. (1998). Analyse des infinitives adnominales en basque. *Lapurdum. Euskal ikerketen aldizkaria | Revue d'études basques | Revista de estudios vascos | Basque studies review*, 3, 37-51. <https://doi.org/10.4000/lapurdum.1669>
- Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. John Benjamins Publishing Company. <https://doi.org/10.1075/sibil.18>
- Paradis, M. (2008). Language and Communication Disorders in Multilinguals. En B. Stemmer & H. A. Whitaker (Eds.), *Handbook of the Neuroscience of Language* (pp. 341-349). Elsevier. <https://doi.org/10.1016/B978-008045352-1.00033-1>
- Parafita Couto, M. C., Munarriz, A., Epelde, I., Deuchar, M., & Oyharçabal, B. (2016). Gender conflict resolution in Spanish–Basque mixed DPs*. *Bilingualism: Language and Cognition*, 19(4), 834-853. <https://doi.org/10.1017/S1366728916000572>
- Pascual y Cabo, D., & Rothman, J. (2012). The (Il)Logical Problem of Heritage Speaker Bilingualism and Incomplete Acquisition. *Applied Linguistics*, 33(4), 450-455. <https://doi.org/10.1093/applin/ams037>
- Pato, E. (2010). La recategorización del adverbio medio en español. *Boletín de filología*, 45(2), 91-110. <https://doi.org/10.4067/S0718-93032010000200005>

- Pearson, B. Z., Fernández, S. C., & Oller, D. K. (1993). Lexical Development in Bilingual Infants and Toddlers: Comparison to Monolingual Norms. *Language Learning*, 43(1), 93-120. <https://doi.org/10.1111/j.1467-1770.1993.tb00174.x>
- Pérez Lagos, M. F. (1986). Composición del verbo más nombre en el DRAE. *LEA: Lingüística Española Actual*, 8(1), 21-58.
- Pérez-Leroux, A. T., & Licerias, J. M. (Eds.). (2002). *The Acquisition of Spanish Morphosyntax* (Vol. 31). Springer. <https://doi.org/10.1007/978-94-010-0291-2>
- Pérez-Pereira, M. (1991). The acquisition of gender: What Spanish children tell us. *Journal of Child Language*, 18(3), 571-590. <https://doi.org/10.1017/S0305000900011259>
- Pérez-Tattam, R., Ezeizabarrena, M. J., Stadthagen-González, H., & Gathercole, V. C. M. (2019). Gender Assignment to Spanish Pseudowords by Monolingual and Basque-Spanish Bilingual Children. *Languages*, 4(3), 58.
- Picallo, M. C. (2008). Gender and number in Romance. *Lingue e linguaggio*, 7(1), 47-66.
- Picone, M. D. (1996). *Anglicisms, neologisms and dynamic French*. Benjamins.
- Piera, C. (1995). On compounding in English and Spanish. En H. Campos & P. M. Kempchinsky (Eds.), *Evolution and revolution in linguistic theory* (pp. 302-314). Georgetown University Press.
- Plag, I. (2006). The variability of compound stress in English: Structural, semantic, and analogical factors. *English Language & Linguistics*, 10(1), 143-172. <https://doi.org/10.1017/S1360674306001821>
- Platzack, C. (1996). The initial hypothesis of syntax: A minimalist perspective on language acquisition and attrition. En H. Clahsen (Ed.), *Generative Perspectives on Language Acquisition. Empirical findings, theoretical considerations and crosslinguistic*

comparisons (pp. 369-368). John Benjamins Publishing Company.

<https://doi.org/10.1075/lald.14.15pla>

Pomerleau, J. (2000). *La adquisición del español en el aula: Los compuestos nominales de los franceses de Canadá* [Tesis de MA]. Universitat de Barcelona.

Poplack, S., Pousada, A., & Sankoff, D. (1982). Competing influences on gender assignment: Variable process, stable outcome. *Lingua*, 57, 1-28.

Porquier, R. (1994). Communication exolingue et contextes d'appropriation. *Bulletin VALS-ASLA*, 59, 159-169.

Portolés, L., & Safont, P. (2018). Examining authentic and elicited data from a multilingual perspective. The real picture of child requestive behaviour in the L3 classroom. *System*, 75, 81-92. <https://doi.org/10.1016/j.system.2018.03.012>

Pratt, C., & Grieve, R. (1984). The development of metalinguistic awareness: An introduction. En W. Tunmer, C. Pratt, & M. L. Herriman (Eds.), *Metalinguistic awareness in children* (pp. 2-11). Springer. https://doi.org/10.1007/978-3-642-69113-3_1

Preston, L. B. (2015). *Generativism and Emergentism: Evidence From Second Language Acquisition Studies of Poverty of the Stimulus Phenomena* [Tesis doctoral, University of Washington]. <https://digital.lib.washington.edu/443/researchworks/handle/1773/33193>

Prévost, P., & White, L. (2000). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second language research*, 16(2), 103-133.

Prodeau, M. (2005). Gender and Number in French L2: Can We Find Out More About the Constraints on Production in L2? En J.-M. Dewaele (Ed.), *Focus on French as a Foreign Language Multidisciplinary Approaches*. Multilingual Matters.

<https://doi.org/10.21832/9781853597688-008>

- Puig-Mayenco, E., González Alonso, J., & Rothman, J. (2020). A systematic review of transfer studies in third language acquisition. *Second Language Research*, 36(1), 31-64.
<https://doi.org/10.1177/0267658318809147>
- Rainer, F. (2016). Spanish. En P. O. Müller, I. Ohnheiser, S. Olsen, & F. Rainer (Eds.), *Word-Formation*. De Gruyter. <https://doi.org/10.1515/9783110379082-012>
- Rainer, F., & Varela, S. (1992). Compounding in Spanish. *Rivista di Linguistica*, 4(1), 117-142.
- Ramsey, M. M. (1956). *A textbook of modern Spanish*. Holt.
- Rao, R. (2015). On the phonological status of Spanish compound words. *Word Structure*, 8(1), 84-118. <https://doi.org/10.3366/word.2015.0074>
- Real Academia Española [RAE]. (1914). *Diccionario de la lengua castellana por la Real Academia Española* (14.^a ed.). Imprenta de los sucesores de Hernando.
- Real Academia Española [RAE]. (1925). *Diccionario de la lengua española* (15.^a ed.). Calpe.
- Real Academia Española [RAE]. (1984). *Diccionario de la lengua española* (20.^a ed.). Espasa-Calpe.
- Real Academia Española [RAE]. (1992). *Diccionario de la lengua española* (21.^a ed.). Espasa-Calpe.
- Real Academia Española [RAE]. (2001a). *Diccionario de la lengua española* (22.^a ed.). Espasa-Calpe.
- Real Academia Española [RAE]. (2001b). *Nuevo tesoro lexicográfico de la lengua española*.
<http://ntlle.rae.es/ntlle/SrvltGUILoginNtlle>
- Real Academia Española [RAE]. (2014). *Diccionario de la lengua española* (23.^a ed.). Espasa.
<http://www.rae.es/rae.html>
- Real Academia Española [RAE]. (2019). *Muestra de novedades DLE 23.3*.
https://www.rae.es/sites/default/files/NOVEDADES_DLE_23.3-Seleccion.pdf

- Real Academia Española [RAE] & Asociación de Academias de la Lengua [ASALE]. (2009). *Nueva gramática de la lengua española*. Espasa Libros.
- Real Academia Española [RAE] & Asociación de Academias de la Lengua [ASALE]. (2011). *Nueva gramática básica de la lengua española*. Espasa.
https://enclave.rae.es/gramatica/html/NG_BASICA_MASTER2AULA.html
- Real Academia Española [RAE], & Asociación de Academias de la Lengua Española [ASALE]. (2010a). La composición. En *Nueva gramática de la lengua española. Manual*. Espasa.
- Real Academia Española [RAE], & Asociación de Academias de la Lengua Española [ASALE]. (2010b). *Nueva gramática de la lengua española. Manual*. Espasa.
- Remysen, W. (2003). *Le français au Québec: Au-delà des mythes*. 15.
- Renner, V., & Fernández-Domínguez, J. (2011). Coordinate compounding in English and Spanish. *Poznań Studies in Contemporary Linguistics*, 47(4), 873-883.
<https://doi.org/10.2478/psicl-2011-0041>
- Restorick Elordi, A. K. (2012). *Interpretation of the French and Spanish Subjunctive by L1, L2, and L3 Speakers: Contexts Where Mood Can Alternate without Ungrammaticality* [Tesis doctoral, University of Western Ontario]. <https://ir.lib.uwo.ca/etd/1055/>
- Rice, C. (2006). Optimizing gender. *Lingua*, 116(9), 1394-1417.
<https://doi.org/10.1016/j.lingua.2004.06.013>
- Riegel, M. (1988a). Les séquences composées N1-N2: Une catégorie floue. *Studia romanica Posnaniensia*, 13(1), 129-138.
- Riegel, M. (1988b). Vrais et faux noms composés: Les séquences binominales en français moderne. En D. Weil & H. Fugier (Eds.), *Actes du Troisième Colloque Régional de Linguistique, Strasbourg, 28-29 avril 1988* (pp. 371-394). Université des Sciences Humaines.

- Ringbom, H. (1987). *The role of the first language in foreign language learning*. Multilingual Matters.
- Rizzi, L. (1982). *Issues in Italian syntax*. Foris Publications.
- Roca de Larios, J., & Manchón Ruiz, R. M. (2006). Algunas consideraciones sobre el factor edad en relación con la enseñanza de las lenguas extranjeras en la escuela. *Porta Linguarum*, 5, 63-76.
- Roca, I. M. (1989). The organisation of grammatical gender. *Transactions of the Philological Society*, 87(1), 1-32.
- Roeper, T., Snyder, W., & Hiramatsu, K. (2002). Learnability in a minimalist framework: Root compounds, merger, and the syntax-morphology interface. *The process of language acquisition*, 25-35.
- Rosenberg, M. (2007). Classification, headedness and pluralization: Corpus evidence from French compounds. *Acta Linguistica Hungarica*, 54(3), 341-360.
<https://doi.org/10.1556/ALing.54.2007.3.4>
- Rosenberg, M. (2008). *La formation agentive en français: Les composés [VN/A/Adv/P]N/A et les dérivés V-ant, V-eur et V-oir(e)* [Tesis doctoral]. Stockholm University.
- Rossovová, L. (2007). *Las palabras compuestas en español* [Tesis de MA]. Masarykova univerzita.
- Rothman, J. (2010). On the typological economy of syntactic transfer: Word order and relative clause high/low attachment preference in L3 Brazilian Portuguese. *IRAL - International Review of Applied Linguistics in Language Teaching*, 48(2-3), 245-273.
<https://doi.org/10.1515/iral.2010.011>

- Rothman, J. (2011). L3 syntactic transfer selectivity and typological determinacy: The typological primacy model. *Second Language Research*, 27(1), 107-127.
<https://doi.org/10.1177/0267658310386439>
- Rothman, J. (2015). Linguistic and cognitive motivations for the Typological Primacy Model (TPM) of third language (L3) transfer: Timing of acquisition and proficiency considered. *Bilingualism: Language and Cognition*, 18(2), 179-190.
<https://doi.org/10.1017/S136672891300059X>
- Rothman, J., & Cabrelli Amaro, J. (2010). What variables condition syntactic transfer? A look at the L3 initial state. *Second Language Research*, 26(2), 189-218.
<https://doi.org/10.1177/0267658309349410>
- Rothman, J., & Slabakova, R. (2018). The generative approach to SLA and its place in modern second language studies. *Studies in Second Language Acquisition*, 40(2), 417-442.
<https://doi.org/10.1017/S0272263117000134>
- Ryder, M. E. (1994). *Ordered Chaos: The Interpretation of English Noun-Noun Compounds*. University of California Press.
- Sabourin, L., Stowe, L. A., & de Haan, G. J. (2006). Transfer effects in learning a second language grammatical gender system. *Second Language Research*, 22(1), 1-29.
<https://doi.org/10.1191/0267658306sr259oa>
- Sagarra, N., & Herschensohn, J. (2010). The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. *Lingua*, 120(8), 2022-2039.
<https://doi.org/10.1016/j.lingua.2010.02.004>
- Salomaa-Robertson, T. (2000). *L2 acquisition of Spanish compounds by native speakers of Finnish*. [Tesis de MA]. University of Ottawa.

- Sanz, C., Park, H. I., & Lado, B. (2015). A functional approach to cross-linguistic influence in *ab initio* L3 acquisition. *Bilingualism: Language and Cognition*, 18(2), 236-251.
<https://doi.org/10.1017/S1366728914000285>
- Scalise, S. (1992). Compounding in Italian. *Italian Journal of Linguistics / Rivista di Linguistica*, 4(1), 175-199.
- Scalise, S., & Bisetto, A. (2011). The Classification of Compounds. En R. Lieber & P. Štekauer (Eds.), *The Oxford handbook of compounding* (pp. 35-53). Oxford University Press.
<https://doi.org/10.1093/oxfordhb/9780199695720.013.0003>
- Scalise, S., & Fábregas, A. (2010). The head in compounding. En S. Scalise & I. Vogel (Eds.), *Cross-disciplinary issues in compounding*. John Benjamins Publishing Company.
- Schmid, M. S. (2013). First language attrition. *WIREs Cognitive Science*, 4(2), 117-123.
<https://doi.org/10.1002/wcs.1218>
- Schmidt, R. (1984). The strengths and limitations of acquisition: A case study of an untutored language learner. *Language, Learning, and Communication*, 3, 1-16.
- Schnell de Acedo, B. (1994). Early morphological development: The acquisition of articles in Spanish. En J. L. Sokolov & C. E. Snow (Eds.), *Handbook of research in language development using CHILDES* (pp. 210-253). Lawrence Erlbaum Associates Publishers.
- Schwartz, B. D., & Sprouse, R. A. (1996). L2 cognitive states and the full transfer/full access model. *Second language research*, 12(1), 40-72.
<https://doi.org/10.1177/026765839601200103>
- Sebastián, E., Soto, P., & Gathercole, V. C. M. (2004). La morfología verbal temprana en español. *Anuario de Psicología*, 35, 18.

- Séguin, H. (1969). *Les marques du genre dans le lexique français écrit contemporain: Compilation des cas et essai de classement. Mémoire de DES* [Tesis doctoral]. Université de Montréal.
- Séguin, Hubert. (1973). Le genre des adjectifs en français: Analyse quantitative et correspondances phonographiques des règles. *Langue française*, 20(1), 52-74.
<https://doi.org/10.3406/lfr.1973.5654>
- Seidler, S. (1988). *Untersuchung zum erwerb von wortbildungsregeln deverbativa nomina (agens und instrument) im französischen* [Tesis de MA]. Universität Hamburg.
- Selinker, L. (1972). Interlanguage. *IRAL - International Review of Applied Linguistics in Language Teaching*, 10(1-4), 209-232. <https://doi.org/10.1515/iral.1972.10.1-4.209>
- Selinker, L. (1983). Language Transfer. En S. M. Gass & L. Selinker (Eds.), *Language transfer in language learning* (Vol. 54, pp. 33-53).
- Selinker, L., & Lamendella, J. T. (1978). Two perspectives on fossilization in interlanguage learning. *Interlanguage Studies Bulletin*, 3(2), 143-191. <https://doi.org/10.2307/43136004>
- Serrano Dolader, D. (2018). La composición en el aula de ELE/L2. En *Formación de palabras y enseñanza del español como LE/L2* (pp. 310-331). Routledge.
- Shoemaker, E., & Rast, R. (2013). Extracting words from the speech stream at first exposure. *Second Language Research*, 29(2), 165-183. <https://doi.org/10.1177/0267658313479360>
- Silva-Corvalán, C., & Treffers-Daller, J. (Eds.). (2016). *Language Dominance in Bilinguals: Issues of Measurement and Operationalization*. Cambridge University Press.
<https://doi.org/10.1017/CBO9781107375345>
- Sisson, C. (2006). *Adult L2 Acquisition of French Grammatical Gender: Investigating sensitivity to phonological and morphological gender cues* [Honours Thesis]. McGill University.

Skinner, B. F. (1957). *Verbal behavior*. Appleton-Century-Crofts. <https://doi.org/10.1037/11256-000>

Slabakova, R. (2002). The compounding parameter in second language acquisition. *Studies in Second Language Acquisition*, 24(04), 507-540.

Slobin, D. I., & Welsh, C. A. (1967). Elicited imitation as a research tool in developmental psycholinguistics. *Working Paper. Language-Behavior Research Laboratory, University of California at Berkeley*, 10, 2-19.

Snyder, W. (1995). *Language acquisition and language variation: The role of morphology* [Tesis doctoral, Massachusetts Institute of Technology]. <http://mitwpl.mit.edu/catalog/snyd01/>

Snyder, W. (2001). On the Nature of Syntactic Variation: Evidence from Complex Predicates and Complex Word-Formation. *Language*, 77(2), 324-342.

<https://doi.org/10.1353/lan.2001.0108>

Snyder, W. (2012). Parameter theory and motion predicates. En V. Demonte & L. McNally (Eds.), *Telicity, Change, and State: A Cross-Categorial View of Event Structure* (pp. 279-299). Oxford University Press.

<https://doi.org/10.1093/acprof:oso/9780199693498.001.0001>

Snyder, W. (2016a). Compound Word Formation. En J. L. Lidz, W. Snyder, & J. Pater (Eds.), *The Oxford Handbook of Developmental Linguistics*.

<https://doi.org/10.1093/oxfordhb/9780199601264.013.6>

Snyder, W. (2016b). How to Set the Compounding Parameter. En L. Perkins, R. Dudley, J. Gerard, & K. Hitczenko (Eds.), *Proceedings of the 6th Conference on Generative Approaches to Language Acquisition North America (GALANA 2015)* (pp. 122-130). Cascadilla Proceedings Project.

- Socarrás, G. M. (2012). *First language acquisition in Spanish: A minimalist approach to nominal agreement*. Continuum International Publishing Group.
- Sorace, A. (2011). Pinning down the concept of “interface” in bilingualism. *Linguistic Approaches to Bilingualism*, 1(1), 1-33. <https://doi.org/10.1075/lab.1.1.01sor>
- Sorace, A., & Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian. *Second Language Research*, 22(3), 339-368. <https://doi.org/10.1191/0267658306sr271oa>
- Sourdôt, M. (1977). Identification et différenciation des unités: Les modalités nominales. En F. François, D. François, E. Sableau-Jouannet, & M. Sourdôt (Eds.), *La syntaxe de l'enfant avant* (Vol. 5, pp. 90-120).
- Starreveld, P. A., De Groot, A. M. B., Rossmark, B. M. M., & Van Hell, J. G. (2014). Parallel language activation during word processing in bilinguals: Evidence from word production in sentence context. *Bilingualism: Language and Cognition*, 17(2), 258-276. <https://doi.org/10.1017/S1366728913000308>
- Steele, S. (1978). Word order variation: A typology study. En J. H. Greenberg, C. A. Ferguson, & E. A. Moravcsik (Eds.), *Universals of Human Language, IV: Syntax* (pp. 585-623). Stanford University Press.
- Suppes, P., Smith, R., & Leveillé, M. (1973). *The French syntax of a child's noun phrases*. Imprimé par Medecine et Hygiène.
- Surridge, M. E. (1989). Le facteur sémantique dans l'attribution du genre aux inanimés en français. *Canadian Journal of Linguistics/Revue Canadienne de Linguistique*, 34(1), 19-44. <https://doi.org/10.1017/S0008413100025871>
- Talmy, L. (1988). The relation of grammar to cognition. En B. Rudzka (Ed.), *Topics in Cognitive Linguistics*. John Benjamins Publishing Company. <https://doi.org/10.1075/cilt.50.08tal>

- Tenorio Rebolledo, I. C. (2005). Las formas de composición nominal en español y sus denominaciones. *Lenguaje*, 33, 379-404. <https://doi.org/10.25100/lenguaje.v33i0.4819>
- Teschner, R. V. (1987). The Gender of French and Spanish Noun Cognates: Some Statistics and a List. *Canadian Modern Language Review/ La Revue canadienne des langues vivantes*, 43(2), 256-266.
- Teschner, R. V., & Russell, W. M. (1984). The gender patterns of Spanish nouns: An inverse dictionary-based analysis. *Hispanic Linguistics*, 1(1), 115-132.
- Thomas, J. (1988). The role played by metalinguistic awareness in second and third language learning. *Journal of Multilingual and Multicultural Development*, 9(3), 235-246. <https://doi.org/10.1080/01434632.1988.9994334>
- Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition* (1.^a ed.). Harvard University Press.
- Tomasello, M. (2009). The usage-based theory of language acquisition. En E. L. Bavin (Ed.), *The Cambridge Handbook of Child Language* (pp. 69-88). Cambridge University Press.
- Toquero Pérez, L. M. (2018). *The Semantics of Spanish Compounding: An Analysis of NN Compounds in the Parallel Architecture* [Tesis de MA, West Virginia University]. <https://doi.org/10.33915/etd.7138>
- Tracy-Ventura, N., McManus, K., Norris, J., & Ortega, L. (2014). Repeat as much as you can': Elicited imitation as a measure of oral proficiency in L2 French. En P. Leclercq, A. Edmonds, & H. Hilton (Eds.), *Measuring L2 proficiency: Perspectives from SLA* (pp. 143-166). Multilingual Matters.
- Trask, R. L. (1995). Origin and relatives of the Basque Language: Review of the evidence. En J. I. Hualde, J. Lakarra, & R. L. Trask (Eds.), *Towards a history of the Basque language*. John Benjamins Publishing Company.

- Trask, R. L. (2003). The Noun Phrase. En J. I. Hualde & J. Ortiz de Urbina (Eds.), *A grammar of Basque* (pp. 113-170). Mouton de Gruyter.
- Treffers-Daller, J. (2019). What Defines Language Dominance in Bilinguals? *Annual Review of Linguistics*, 5(1), 375-393. <https://doi.org/10.1146/annurev-linguistics-011817-045554>
- Trías, M., & Villanueva, A. M. (2011). La composición nominal en la adquisición del español como segunda lengua. Respuestas a una prueba de evocación. *Entre Lenguas*, 16, 63-78.
- Trías, M., & Villanueva, A. M. (2013). El Parámetro de los Compuestos en la adquisición del español como lengua extranjera considerando el tipo de lengua materna y el contexto de aprendizaje. *Colombian Applied Linguistics Journal*, 15(1), 58-75.
- Tucker, G. R., Lambert, W. E., & Rigault, A. (1977). *The French speaker's skill with grammatical gender: An example of rule-governed behavior*. Mouton.
- Tucker, G. R., Lambert, W. E., Rigault, A., & Segalowitz, N. (1968). A psychological investigation of French speakers' skill with grammatical gender. *Journal of Verbal Learning and Verbal Behavior*, 7(2), 312-316. [https://doi.org/10.1016/S0022-5371\(68\)80007-6](https://doi.org/10.1016/S0022-5371(68)80007-6)
- Tuggy, D. (2003). *Abrelatas And Scarecrow Nouns: Exocentric Verb-Noun Compounds As Illustrations of Basic Principles of Cognitive Grammar*. 3(2), 25-51.
- Uhlenbeck, C. C. (1928). Quelques observations sur le mot ilargi. En *Homenaje a d. Carmelo de Echeagaray* (pp. 557-560).
- Unsworth, S. (2016). Amount of exposure as a proxy for dominance in bilingual language acquisition. En C. Silva-Corvalan & J. Treffers-Daller (Eds.), *Language Dominance in Bilinguals* (pp. 156-173). Cambridge University Press.
- <https://doi.org/10.1017/CBO9781107375345.008>

- Vainikka, A., & Young-Scholten, M. (1996a). Gradual development of L2 phrase structure. *Second language research*, 12(1), 7-39. <https://doi.org/10.1177/026765839601200102>
- Vainikka, A., & Young-Scholten, M. (1996b). The early stages in adult L2 syntax: Additional evidence from Romance speakers. *Second Language Research*, 12(2), 140-176. <https://doi.org/10.1177/026765839601200202>
- Val Álvaro, J. F. (1999). La composición. En I. Bosque & V. Demonte (Eds.), *Gramática descriptiva de la lengua española* (Vol. 3, pp. 4757-4840). Espasa.
- Valencia, J. F., & Cenoz, J. (1992). The Role of Bilingualism in Foreign Language Acquisition: Learning English in the Basque Country. *Journal of Multilingual and Multicultural Development*, 13(5), 433-449.
- Valenzuela, E., Faure, A., Ramírez-Trujillo, A. P., Barski, E., Pangtay, Y., & Diez, A. (2012). Gender and Heritage Spanish Bilingual Grammars: A Study of Code-mixed Determiner Phrases and Copula Constructions. *Hispania*, 95(3), 481-494. JSTOR.
- Van Valin, R. D., & LaPolla, R. J. (1997). *Syntax structure, meaning and function*. Cambridge University Press.
- Varela, S. (1989). Spanish endocentric compounds and the 'atom condition'. En C. Kirschner & J. A. DeCesaris (Eds.), *Current Issues in Linguistic Theory* (Vol. 60, p. 397). John Benjamins Publishing Company. <https://doi.org/10.1075/cilt.60.24var>
- Varela, S. (2012). Derivation and Compounding. En *The Handbook of Hispanic Linguistics* (pp. 209-226). Blackwell Publishing. <https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118228098.ch11>
- Villoing, F. (2003). Les mots composés VN du français: Arguments en faveur d'une construction morphologique. *Cahiers de Grammaire*, 28, 183-196.

- Villoing, F. (2009). Les mots composés VN. En B. Fradin, F. Kerleroux, & M. Plénat (Eds.), *Aperçus de morphologie du français* (pp. 175-198). Presses universitaires de Vincennes.
- Villoing, F. (2012). French compounds. *Probus*, 24(1), 29-60. <https://doi.org/10.1515/probus-2012-0003>
- Volterra, V., & Taeschner, T. (1978). The acquisition and development of language by bilingual children. *Journal of child language*, 5(2), 311-326.
<https://doi.org/10.1017/S0305000900007492>
- von Heusinger, K., & Schwarze, C. (2013). Italian V+N compounds, inflectional features and conceptual structure. *Morphology*, 23(3), 325-350. <https://doi.org/10.1007/s11525-013-9210-z>
- Weinreich, U. (1979). *Languages in contact: Findings and problems*. Mouton. (Original work published 1953)
- Westergaard, M. (2009a). *The Acquisition of Word Order: Micro-cues, information structure, and economy*. John Benjamins Publishing Company. <https://doi.org/10.1075/la.145>
- Westergaard, M. (2009b). Usage-based vs. rule-based learning: The acquisition of word order in wh-questions in English and Norwegian. *Journal of Child Language*, 36(05), 1023-1051.
- Westergaard, M. (2014). Linguistic variation and micro-cues in first language acquisition. *Linguistic Variation*, 14(1), 26-45. <https://doi.org/10.1075/lv.14.1.02wes>
- Westergaard, M. (2019). Microvariation in multilingual situations: The importance of property-by-property acquisition: *Second Language Research*. Online first.
<https://doi.org/10.1177/0267658319884116>
- Westergaard, M., Mitrofanova, N., Mykhaylyk, R., & Rodina, Y. (2017). Crosslinguistic influence in the acquisition of a third language: The Linguistic Proximity Model.

International Journal of Bilingualism, 21(6), 666-682.

<https://doi.org/10.1177/1367006916648859>

White, L. (2003). *Second language acquisition and universal grammar*. Cambridge University Press.

White, L., Valenzuela, E., Kozłowska–Macgregor, M., & Leung, Y.-K. I. (2004). Gender and number agreement in nonnative Spanish. *Applied Psycholinguistics*, 25(01).

<https://doi.org/10.1017/S0142716404001067>

Williams, S., & Hammarberg, B. (1998). Language Switches in L3 Production: Implications for a Polyglot Speaking Model. *Applied Linguistics*, 19(3), 295-333.

<https://doi.org/10.1093/applin/19.3.295>

Yang, J. (2016). *El género gramatical en una gramática del español como segunda lengua* [Tesis doctoral, Universidad Complutense de Madrid]. <http://eprints.ucm.es/40011/1/T37963.pdf>

Yoon, J. (2009). Constructional meanings of verb–noun compounds in Spanish: Limpiabotas vs. tientaparedes. *Language Sciences*, 31(4), 507-530.

<https://doi.org/10.1016/j.langsci.2007.09.005>

Zwicky, A. M. (1985). Heads. *Journal of Linguistics*, 21(1), 1-29.

Zytsar, Y. (2000). Análisis crítico de los desarrollos etimológicos de Uhlenbeck y Michelena para el vasco ilargi “luna”. *Fontes Linguae Vasconum*, 83, 29-36.

Zyzik, E. (2017). Subject expression in L2 Spanish: Convergence of generative and usage-based perspectives? *Second Language Research*, 33(1), 33-59.

<https://doi.org/10.1177/0267658316659106>

Zyzik, E. (2019). Incomplete acquisition from a usage-based perspective: A response to Domínguez, Hicks, and Slabakova. *Studies in Second Language Acquisition*, 41(2), 279-282. <https://doi.org/10.1017/S0272263119000330>

Anexo 1. **Cuestionario para los padres (versión francesa)**

Questionnaire sur le profil linguistique de votre enfant

Ce court questionnaire doit être rempli par l'un des parents de l'élève. Veuillez prendre le temps de le remplir de la manière la plus précise possible. Certaines questions vous concernent, mais la plupart des questions concernent votre enfant. Vous êtes sans doute la personne la mieux placée pour répondre, mais en cas de doute, n'hésitez pas à demander l'avis de votre enfant.

Questions sur le(s) parent(s)

Les questions de cette section ne concernent que le(s) parent(s) de l'enfant.

Votre prénom

Votre nom de famille

Sexe

Masculin

Féminin

Votre courriel

Votre numéro de téléphone

Pays (et région) d'origine du père

Langue première (et toujours dominante) du père

Pays (et région) d'origine de la mère

Langue première (et toujours dominante) de la mère

Suivant ▶

Questionnaire sur le profil linguistique de votre enfant

À partir de cette étape, les informations demandées portent sur votre enfant.

Prénom

Nom

Sexe

Masculin

Féminin

Date de naissance

Lieu de naissance (Pays, région et ville)

Date d'arrivée en Iparralde, si né(e) à l'extérieur

Année scolaire actuelle

Âge d'entrée à l'Ikastola

École

Enseignant(e)

Langue principalement parlée avec le père

Langue principalement parlée avec la mère

Avez-vous autre chose à signaler à propos des langues parlées à la maison?

◀ Précédent

Suivant ▶

Questionnaire sur le profil linguistique de votre enfant

Questions spécifiques sur les différentes langues apprises par votre enfant
(langue **1** / 3)

Veuillez remplir cette section en fonction de l'**ordre chronologique** d'acquisition/apprentissage des langues de votre enfant. Par exemple, s'il a appris le français d'abord à la maison, le basque à l'ikastola et l'espagnol au collège, vous devriez respectivement choisir le français comme première langue, le basque comme deuxième et l'espagnol en troisième. Si votre enfant a appris simultanément deux langues dès sa naissance, vous pouvez choisir comme première langue la langue qu'il maîtrise le mieux aujourd'hui.

Langue

Euskara

Âge de début d'acquisition de cette langue

Saisir 0 si votre enfant a commencé à apprendre cette langue dès sa naissance

Où a-t-il principalement appris cette langue?

- À la maison
- À la garderie / pré-maternelle
- À l'école primaire
- Au collège
- Autre

Avec qui a-t-il principalement appris cette langue?

- Mère
- Père
- Les deux parents
- Enseignant(e)
- Autre

À votre avis, comment est le niveau de votre enfant pour cette langue?

Excellent

Pendant combien d'heures par semaine utilise-t-il ou est-il exposé à cette langue?

Entrez une valeur approximative

Où?

Avec qui?

◀ Précédent

Suivant ▶

Questionnaire sur le profil linguistique de votre enfant

Questions spécifiques sur les différentes langues apprises par votre enfant
(langue **2** / 3)

Si votre enfant ne connaît plus d'autres langues, vous pouvez passer à l'étape suivante directement avec le bouton rouge qui apparaît sous ce paragraphe. Toutefois, même s'il ne possède que des connaissances élémentaires d'une autre langue, merci de bien nous l'indiquer.

Passer 

Langue

Français

Âge de début d'acquisition de cette langue

Saisir 0 si votre enfant a commencé à apprendre cette langue dès sa naissance

Où a-t-il principalement appris cette langue?

- À la maison
- À la garderie / pré-maternelle
- À l'école primaire
- Au collège
- Autre

Avec qui a-t-il principalement appris cette langue?

- Mère
- Père
- Les deux parents
- Enseignant(e)
- Autre

À votre avis, comment est le niveau de votre enfant pour cette langue?

Excellent

Pendant combien d'heures par semaine utilise-t-il ou est-il exposé à cette langue?

Entrez une valeur approximative

Où?

Avec qui?

◀ Précédent

Suivant ▶

Questionnaire sur le profil linguistique de votre enfant

Questions sur les préférences des différentes langues par votre enfant

Ces questions nous aideront à déterminer quelle est la langue dominante de votre enfant. Vous pouvez lui poser les questions directement.

Langue préférée pour lire

Euskara

Langue préférée pour écrire

Euskara

Langue préférée pour compter

Euskara

Langue préférée pour parler de sujets qui l'intéressent

Euskara

Veuillez nous indiquer toute autre information qui pourrait être pertinente à propos de votre enfant.

◀ Précédent

Envoyer ▶

Anexo 2. Mini-diccionario visual

									
BOSQUE	PAQUETE	PUERTA	CARTA						
									
BOTELLA	UÑA	TELÉFONO	CABLE						
									
BICICLETA	FIESTA	BARCO	COCHE						
									
BOTA	BOLÍGRAFO	SILLA	ESTRELLA						
									
BOMBA	CORAZÓN	SORPRESA	LATA						
									
PLATO	UVA	DIENTE	PATATA						
									
PASTEL	MONSTRUO	MELÓN	PIRATA						
									
POLICÍA	MUJER	HOMBRE	REINA						
									
GATO	PERRO	RANA	PULPO						
			<table border="1"> <tbody> <tr> <td>AMARILLO/A</td> <td>GRIS</td> </tr> <tr> <td>ROJO/A</td> <td>VERDE</td> </tr> <tr> <td>BLANCO/A</td> <td>AZUL</td> </tr> </tbody> </table>	AMARILLO/A	GRIS	ROJO/A	VERDE	BLANCO/A	AZUL
AMARILLO/A	GRIS								
ROJO/A	VERDE								
BLANCO/A	AZUL								
HORMIGA	ARAÑA	ABEJA	COLORES						

Anexo 3. Lista de ítems para elicitación del género

Ítems con un adjetivo variable en español

Español	Francés	Gén. (ES)	Terminación	Analógico
una nariz roja	un nez vert	Fem.	Opaco	No
una nube amarilla	un nuage gris	Fem.	Opaco	No
una pared blanca	un mur blanc	Fem.	Opaco	No
una flor blanca	une fleur blanche	Fem.	Opaco	Sí
una llave amarilla	une clé grise	Fem.	Opaco	Sí
una cruz blanca	une croix blanche	Fem.	Opaco	Sí
una cama roja	un lit vert	Fem.	Trans.	No
una camiseta blanca	un t-shirt blanc	Fem.	Trans.	No
una mochila blanca	un sac blanc	Fem.	Trans.	No
una casa roja	une maison verte	Fem.	Trans.	Sí
una guitarra amarilla	une guitare-violette	Fem.	Trans.	Sí
una manzana roja	une pomme verte	Fem.	Trans.	Sí
un bigote amarillo	une moustache grise	Masc.	Opaco	No
un cohete rojo	une fusée verte	Masc.	Opaco	No
un tomate rojo	une tomate verte	Masc.	Opaco	No
un corazón rojo	un cœur vert	Masc.	Opaco	Sí
un guante blanco	un gant blanc	Masc.	Opaco	Sí
un lápiz amarillo	un crayon violet	Masc.	Opaco	Sí
un armario blanco	une armoire blanche	Masc.	Trans.	No
un plátano amarillo	une banane grise	Masc.	Trans.	No
un vestido blanco	une robe blanche	Masc.	Trans.	No
un castillo amarillo	un château gris	Masc.	Trans.	Sí
un libro rojo	un livre rouge	Masc.	Trans.	Sí
un sombrero amarillo	un chapeau gris	Masc.	Trans.	Sí

Nota. El color de los ítems variaba de una lengua a otra cuando el adjetivo equivalente en francés era invariable.

Ítems con un adjetivo invariable en español

Español	Francés	Gén. (ES)	Terminación	Analógico
una nuez gris	une noix jaune	Fem.	Opaco	Sí
una fuente azul	une fontaine bleue	Fem.	Opaco	Sí
una torre gris	une tour jaune	Fem.	Opaco	Sí
una botella gris	une bouteille jaune	Fem.	Trans.	Sí
una mesa verde	une table rouge	Fem.	Trans.	Sí
una silla verde	une chaise rouge	Fem.	Trans.	Sí
un árbol verde	un arbre rouge	Masc.	Opaco	Sí
un pan gris	un pain jaune	Masc.	Opaco	Sí
un pantalón azul	un pantalon bleu	Masc.	Opaco	Sí
un pepino verde	un concombre vert	Masc.	Trans.	Sí
un barco azul	un bateau bleu	Masc.	Trans.	Sí
un bolígrafo azul	un stylo bleu	Masc.	Trans.	Sí

Nota. El color de los ítems variaba de una lengua a otra cuando el adjetivo equivalente en francés era variable.

Anexo 4. **Lista de ítems para elicitación los compuestos N+N**

ES	FR	EU	Gén.	Atest.	Anim.
bicicleta pirata	vélo pirate	pirata txirindula	F	No	No
silla pirata	chaise pirate	pirata kadira	F	No	No
bolígrafo pirata	stylo pirate	pirata idazluma	M	No	No
paquete sorpresa	paquet surprise	sorpresa pakete	M	No	No
mujer pulpo	femme pieuvre	olagarro emazte	F	No	Sí
mujer hormiga	femme fourmi	xinaurri emazte	F	No	Sí
gato policía	chat policier	polizia katu	M	No	Sí
hombre hormiga	homme fourmi	xinaurri gizon	M	No	Sí
estrellamar	étoile de mer	itsas izar	F	Sí*	No
carta bomba	lettre bombe	bonba gutun	F	Sí	No
barco pirata	bateau pirate	pirata itsasontzi	M	Sí	No
coche bomba	voiture-bombe	bonba auto	M	Sí	No
abeja reina	abeille reine	erle erregina	F	Sí	Sí
mujer araña	femme-araignée	armiarma emazte	F	Sí	Sí
perro policía	chien policier	polizia txakur	M	Sí	Sí
hombre araña	homme-araignée	armiarma gizon	M	Sí	Sí

Nota. Los equivalentes en francés y en euskera son posibles gramaticalmente, lo cual no significa que están atestiguados como sus equivalentes en español. En cualquier caso, los equivalentes que aparecen en la tabla representan formas posibles, pero también aceptamos otras respuestas como correctas.

Anexo 5. **Lista de ítems para elicitación de los compuestos deverbales**

ES	FR	EU	Atest.*	Anim.
un guardabosques	garde-forêt	oihan zaintzaile	Sí	Sí
un limpiabotas	lave-chaussure	zapata garbitzaile	Sí	Sí
un rompecorazones	brise-cœur	bihotz hausle	Sí	Sí
un pisaúvas	foule-raisin	mahats zapaltzaile	Sí	Sí
un abrebotellas	ouvre-bouteille	botila irekigailu	Sí	No
un abrelatas	ouvre-boite	kontserba irekigailu	Sí	No
un lavaplatos	lave-vaisselle	azieta garbigailu	Sí	No
un pelapatatas	épluche patate	lursagar zurigailu	Sí	No
un abrepuertas	ouvre-porte	ate irekitzaile	No	Sí
un arrancadientes	arrache-dent	hortz ateratzaile	No	Sí
un comemelones	mange-melon	meloï jale	No	Sí
un lavabiciquetas	lave-vélo	bizikleta garbitzaile	No	Sí
un atrapamonstruos	attrape-monstre	munstro harrapagailu	No	No
un rompebotellas	brise-bouteille	botila hautsigailu	No	No
un cortacables	coupe-cable	kable mozgailu	No	No
un lanzapasteles	lance-gâteau	bixkotx aurtikigailu	No	No

Nota. Los equivalentes en francés y en euskera son posibles gramaticalmente, lo cual no significa que están atestiguados como sus equivalentes en español. En cualquier caso, los equivalentes que aparecen en la tabla representan formas posibles, pero también aceptamos otras respuestas como correctas.

Anexo 6. **Modelo de la evolución longitudinal de la concordancia de género.**

Regresión logística de efectos mixtos. Los niveles de referencia son Grupo=EU, Género=MASC, Tipo=DET, Transparente=Opaco, Congruente=Incongruente.

	β	SE	z	Pr(> z)
(Intercepto)	0.797	0.191	4.172	< 0.001
Recolección (~140 horas)	0.988	0.082	12.105	< 0.001
Recolección (~200 horas)	1.458	0.085	17.114	< 0.001
Grupo (FR)	-0.442	0.233	-1.896	.058
Grupo (FR/EU)	-0.650	0.207	-3.140	.002
Género (FEM)	-2.396	0.283	-8.455	< 0.001
Tipo (ADJ)	-0.035	0.069	-0.509	.611
Transparente	1.598	0.302	5.298	< 0.001
Congruente	0.649	0.300	2.167	.030
Género (FEM) × Tipo (ADJ)	-1.517	0.139	-10.898	< 0.001
Grupo (FR) × Transparente	-1.070	0.352	-3.042	.002
Grupo (FR/EU) × Transparente	-0.427	0.312	-1.368	.171
Grupo (FR) × Congruente	1.141	0.353	3.237	.001
Grupo (FR/EU) × Congruente	0.365	0.310	1.177	.239

Anexo 7. **Modelo de la concordancia de género (recolección 2)**

Regresión logística de efectos mixtos. Los niveles de referencia son Grupo=EU, Género=MASC, Tipo=DET, Transparente=Opaco, Congruente=Incongruente.

	β	SE	z	Pr(> z)
(Intercepto)	1.936	0.311	6.218	< 0.001
Grupo (FR)	-0.385	0.398	-0.966	.334
Grupo (FR/EU)	-0.706	0.354	-1.992	.046
Género (FEM)	-2.071	0.416	-4.976	< 0.001
Tipo (ADJ)	-0.015	0.116	-0.133	.894
Transparente	1.907	0.491	3.886	< 0.001
Congruente	0.369	0.437	0.844	.399
Género (FEM) × Tipo (ADJ)	-1.199	0.233	-5.155	< 0.001
Grupo (FR) × Transparente	-1.412	0.594	-2.379	.017
Grupo (FR/EU) × Transparente	-0.279	0.527	-0.530	.596
Grupo (FR) × Congruente	1.660	0.526	3.156	.002
Grupo (FR/EU) × Congruente	0.899	0.458	1.962	.049

Anexo 8. **Modelo de la concordancia de género (recolección 3)**

Regresión logística de efectos mixtos. Los niveles de referencia son Grupo=EU, Género=MASC, Tipo=DET, Transparente=Opaco, Congruente=Incongruente.

	β	SE	z	Pr(> z)
(Intercepto)	2.802	0.400	7.012	< 0.001
Grupo (FR)	-0.517	0.502	-1.030	.303
Grupo (FR/EU)	-0.546	0.451	-1.209	.227
Género (FEM)	-2.702	0.492	-5.489	< 0.001
Tipo (ADJ)	-0.093	0.138	-0.680	.497
Transparente	3.024	0.661	4.574	< 0.001
Congruente	0.782	0.462	1.694	.090
Género (FEM) × Tipo (ADJ)	-0.379	0.275	-1.376	.169
Grupo (FR) × Transparente	-1.173	0.801	-1.464	.143
Grupo (FR/EU) × Transparente	0.112	0.727	0.154	.878
Grupo (FR) × Congruente	1.319	0.537	2.457	.014
Grupo (FR/EU) × Congruente	0.313	0.474	0.661	.509

Anexo 9. **Resultados por ítem en la concordancia determinante – sustantivo**

Género	Transp.	Congruente	Ítem	EU	FR	FR/EU
M	No	No	un bigote amarillo	92 %	64 %	61 %
M	No	No	un cohete rojo	73 %	45 %	59 %
M	No	No	un tomate rojo	58 %	64 %	35 %
M	No	Sí	un corazón rojo	70 %	90 %	75 %
M	No	Sí	un guante blanco	100 %	82 %	78 %
M	No	Sí	un lápiz amarillo	100 %	100 %	76 %
M	Sí	No	un armario blanco	83 %	60 %	43 %
M	Sí	No	un plátano amarillo	92 %	55 %	50 %
M	Sí	No	un vestido blanco	73 %	45 %	47 %
M	Sí	Sí	un castillo amarillo	82 %	100 %	72 %
M	Sí	Sí	un libro rojo	83 %	100 %	83 %
M	Sí	Sí	un sombrero amarillo	100 %	82 %	82 %
F	No	No	una nariz roja	11 %	10 %	29 %
F	No	No	una nube amarilla	25 %	18 %	25 %
F	No	No	una pared blanca	17 %	9 %	33 %
F	No	Sí	una cruz blanca	36 %	27 %	39 %
F	No	Sí	una flor blanca	50 %	55 %	50 %
F	No	Sí	una llave amarilla	55 %	73 %	38 %
F	Sí	No	una cama roja	67 %	27 %	25 %
F	Sí	No	una camiseta blanca	67 %	27 %	41 %
F	Sí	No	una mochila blanca	75 %	55 %	44 %
F	Sí	Sí	una casa roja	83 %	91 %	67 %
F	Sí	Sí	una guitarra amarilla	75 %	82 %	53 %
F	Sí	Sí	una manzana roja	73 %	82 %	50 %

Anexo 10. **Resultados por ítem en la concordancia sustantivo - adjetivo**

Género	Transp.	Congruente	Ítem	EU	FR	FR/EU
M	No	No	un bigote amarillo	100 %	62 %	86 %
M	No	No	un cohete rojo	83 %	69 %	81 %
M	No	No	un tomate rojo	92 %	92 %	91 %
M	No	Sí	un corazón rojo	83 %	100 %	100 %
M	No	Sí	un guante blanco	92 %	100 %	86 %
M	No	Sí	un lápiz amarillo	100 %	100 %	95 %
M	Sí	No	un armario blanco	100 %	92 %	80 %
M	Sí	No	un plátano amarillo	100 %	54 %	91 %
M	Sí	No	un vestido blanco	100 %	60 %	88 %
M	Sí	Sí	un castillo amarillo	100 %	92 %	95 %
M	Sí	Sí	un libro rojo	100 %	100 %	89 %
M	Sí	Sí	un sombrero amarillo	100 %	92 %	86 %
F	No	No	una nariz roja	0 %	0 %	7 %
F	No	No	una nube amarilla	9 %	15 %	14 %
F	No	No	una pared blanca	0 %	0 %	5 %
F	No	Sí	una cruz blanca	8 %	0 %	14 %
F	No	Sí	una flor blanca	0 %	31 %	18 %
F	No	Sí	una llave amarilla	17 %	46 %	18 %
F	Sí	No	una cama roja	27 %	9 %	16 %
F	Sí	No	una camiseta blanca	33 %	15 %	9 %
F	Sí	No	una mochila blanca	25 %	15 %	9 %
F	Sí	Sí	una casa roja	18 %	31 %	5 %
F	Sí	Sí	una guitarra amarilla	25 %	50 %	36 %
F	Sí	Sí	una manzana roja	17 %	33 %	14 %

Anexo 11. Certificado ético



N° de certificat
CERAS-2016-17-049-D

Comité d'éthique de la recherche en arts et en sciences

CERTIFICAT D'APPROBATION ÉTHIQUE

Le Comité d'éthique de la recherche en arts et en sciences (CÉRAS), selon les procédures en vigueur, en vertu des documents qui lui ont été fournis, a examiné le projet de recherche suivant et conclu qu'il respecte les règles d'éthique énoncées dans la Politique sur la recherche avec des êtres humains de l'Université de Montréal.

Projet	
Titre du projet	Influences translinguistiques et transferts morphosyntaxiques chez les apprenants d'espagnol langue tierce au Pays basque nord (France) et dans la région de Montréal (Québec)
Étudiant requérant	Hugues Lacroix Étudiant au doctorat, FAS-Département de littératures et de langues du monde
Sous la direction de	Anahi Alba de la Fuente, professeure adjointe, FAS-Département de littératures et langues du monde, Université de Montréal
Financement	
Organisme	FRQSC
Programme	Bourse doctorat en recherche
Titre de l'octroi si différent	
Numéro d'octroi	201873
Chercheur principal	
No de compte	

MODALITÉS D'APPLICATION

Tout changement anticipé au protocole de recherche doit être communiqué au CÉRAS qui en évaluera l'impact au chapitre de l'éthique.

Toute interruption prématurée du projet ou tout incident grave doit être immédiatement signalé au CÉRAS.

Selon les règles universitaires en vigueur, un suivi annuel est minimalement exigé pour maintenir la validité de la présente approbation éthique, et ce, jusqu'à la fin du projet. Le questionnaire de suivi est disponible sur la page web du CÉRAS.


Marin Arguin, Président
Comité d'éthique de la recherche en arts
et en sciences
Université de Montréal

11 juillet 2016
Date de délivrance

31 juillet 2021
Date de fin de Validité

adresse postale
C.P. 6128, succ. Centre-ville
Montréal QC H3C 3J7

adresse civique
Pavillon Lionel-Groulx
3150, rue Jean-Brillant
Local C-9104
Montréal QC H3T 1N8

Téléphone : 514-343-7338
ceras@umontreal.ca
www.ceras.umontreal.ca