

Université de Montréal

Contrainte sur la mémoire immédiate des groupes intonatifs comme principe de
structuration de la prosodie

par
Annie Gilbert

Département de linguistique et de traduction
Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de maîtrise
en linguistique

Avril 2006

© Annie Gilbert 2006



P
25
U54
2006
v.011

AVIS

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

NOTICE

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal
Faculté des études supérieures

Ce mémoire intitulé :

Contrainte sur la mémoire immédiate des groupes intonatifs comme principe de
structuration de la prosodie

présenté par :

Annie Gilbert

a été évalué par un jury composé des personnes suivantes :

..... M. Richard Patry

président-rapporteur

..... M. Victor Boucher

directeur de recherche

..... Mme Brigitte Stemmer

membre du jury

Mémoire accepté le 28 juin 2006

Résumé et mots clés

Le manque de correspondance entre les constituants syntaxiques et les structures prosodiques de la parole constitue un problème fondamental en sciences du langage. Le présent mémoire porte sur le rôle structurant des groupes intonatifs (GI) et l'influence de processus mnémoniques. Une recension sert à clarifier la terminologie, à démontrer l'avantage du français dans l'étude de la prosodie et à préciser la question « des effets de longueur » au centre de certains modèles. Ces effets portent sur des observations d'un nombre limite de syllabes dans les groupes de rythme et d'intonation, ce qui suggère un principe d'organisation indépendant de la syntaxe. On fait état d'un parallèle entre ces observations et des contraintes bien connues sur la mémoire sérielle immédiate. En soi, l'hypothèse d'un lien entre la mémoire et la prosodie de la parole semble logique puisque toute contrainte agissant sur la mémoire et la reproduction de rythmes et de groupes intonatifs limiterait de facto les schèmes pouvant être appris et utilisés en parole. Suivant cette prémisse, trois expériences ont été élaborées afin d'explorer le lien entre les contraintes sur la mémoire de GI et les effets de longueur. L'Expérience I avait pour objectif d'établir des statistiques sur la longueur des GI en parole. Une analyse acoustique de discours spontanés de 15 locuteurs natifs du français a permis d'établir à huit syllabes la longueur limite des GI en parole. L'Expérience II, faite auprès de 31 sujets, portait sur le rappel de séries entendues avec des GI de deux à neuf syllabes. Les principaux résultats montrent que le rappel de schèmes intonatifs varie suivant la longueur des groupes et que les GI de 8 et 9 syllabes sont significativement plus difficiles à rappeler que les GI de 5 et 6 syllabes. L'Expérience III visait à évaluer l'applicabilité de la limite de huit syllabes à la prédiction des GI dans des phrases à syntagmes de longueur variable présentées visuellement. L'analyse acoustique des répétitions de 40 sujets démontre que la contrainte de longueur des GI est un meilleur facteur de prédiction de la prosodie produite que la syntaxe seule. Pris ensemble, les résultats des trois expériences ont des implications sur les modèles de prosodie où on prend pour acquis que les schèmes intonatifs sont déterminés par la syntaxe.

Mots clés : intonation, prosodie, mémoire immédiate, structure prosodique

Abstract and key words

The lack of correspondence between syntactic constituents and speech prosody presents a pivotal problem in language research. This study focuses on the structural functions of intonational groups (IGs). The literature review clarifies the terminology, the specific advantage of French in observing prosodic structures, and the issue of length effects in models of prosody. By these effects, when phrases exceed a given number of syllables, a change occurs in rhythm and tone groups, which as such suggests a principle of prosodic organization that is independent of syntax. No study has explained these effects yet. However, the review outlines an interesting parallel between reported length restrictions on prosodic structures and constraints on short-term serial memory. As such, the link between memory and prosodic structure is a logical one: any constraint on memory and the reproducibility of prosodic patterns would as such limit the learning and use of rhythm and tone groups to a given set of patterns. On this basis, the present study examines the link between memory constraints and length effects on IGs by way of three experiments. In Experiment I, acoustic analyses served to gather descriptive statistics on the size of IGs in the spontaneous speech of fifteen speakers of French. The results reveal an upper-limit of about eight syll. per IG, included in two rhythm groups. Experiment II involved a task where 31 subjects (excluding subjects with musical training) were asked to recall stimuli constructed with IGs of 2 to 9 syll. The main results show that correct recall of the number of syll. in IGs (or correct location of an IG boundary) varies proportionally with the size of the groups. There is no sharp limit, though recall of IGs of 8-9 syll. is significantly weaker than recall of IGs of 5-6 syll. and varying rhythm groups did not alter recall performance. Experiment III was aimed at evaluating whether an 8-syll. limit predicts IGs of given sentences where verb-phrases (VP) vary in numbers of syll. Visually presented sentences were repeated by 31 subjects. The results of acoustic analyses show that the syllable-count principle provides a better predictor of IGs than syntactic structures. Taken together, the results bear several implications on current views and models, which assume that prosody derives from syntax.

Key words : intonation; prosody, short-term memory; prosodic structure.

Table des matières

Liste des tableaux	vii
Liste des figures	viii
Liste des sigles et abréviations	x
Remerciements.....	xi
1. Introduction : La prosodie dans son ensemble.....	1
1.1. Types de prosodie.....	1
1.2. Constituants de la prosodie phonologique et phonétique.....	1
1.3. Avantage du français dans l'étude de la prosodie structurante.....	5
1.4. Un problème fondamental : l'incongruité entre la syntaxe et la prosodie.....	5
1.4.1. Incongruité lors de l'acquisition.....	6
1.4.2. Effets de longueur.....	6
1.5. Fonctions de la prosodie structurante.....	7
1.6. Conclusion et objectifs.....	7
2. Recension de la littérature.....	9
2.1. Clarification de la terminologie mnémonique.....	9
2.2. Clarification de la terminologie et des corrélats observables des constituants prosodiques.....	10
2.2.1. La superposition des paramètres acoustiques : le cas de l'anglais.....	10
2.2.2. La superposition des paramètres acoustiques : le cas (facilitant) du français.....	12
2.3. De la syllabe comme unité de base.....	14
2.4. Le groupe rythmique.....	15
2.4.1. Le groupe rythmique et ses corrélats phonétiques.....	15
2.4.2. Sur le rôle structurant du groupe rythmique.....	17
2.5. Le groupe intonatif.....	18
2.5.1. Facteurs confondant l'observation des groupes intonatifs.....	18
2.5.2. Le groupe intonatif et ses corrélats phonétiques.....	18
2.5.3. Contenu des groupes intonatifs.....	19
2.5.4. Sur le rôle structurant du groupe intonatif.....	20
2.6. Les précédents : de la fonction structurante de la prosodie et ses liens avec la mémoire immédiate.....	21
2.7. Conclusions : le rôle structurant des groupes intonatifs et les leurs liens avec les processus de mémorisation.....	23
3. Hypothèses.....	26
4. Expérience I.....	28
4.1. Sujets.....	28

4.2. Procédure.....	29
4.3. Analyse.....	29
4.4. Résultats.....	32
4.5. Discussion.....	33
5. Expérience II.....	34
5.1. Stimuli.....	37
5.2. Sujets.....	40
5.3. Procédure.....	40
5.4. Analyse.....	42
5.5. Résultats.....	43
5.6. Discussion.....	46
6. Expérience III.....	49
6.1. Sujets.....	49
6.2. Stimuli.....	50
6.3. Procédure.....	52
6.4. Analyse.....	53
6.5. Résultats.....	53
6.6. Discussion.....	55
7. Discussion générale et conclusions.....	56
8. Bibliographie.....	59
Annexe.....	xii

Liste des tableaux

Tableau I : *Stimuli utilisés dans l'Expérience II*

Tableau II : *Test non paramétrique de Friedman comparant l'ensemble des GI en position initiale de séquence.*

Tableau III : *Statistiques descriptives en fonction de la longueur du GI en position initiale.*

Tableau IV : *Phrases du premier groupe en fonction du nombre de syllabes avant et après la frontière syntaxique majeure.*

Tableau V : *Phrases du deuxième groupe en fonction du nombre de syllabes dans les constituants*

Tableau VI : *Nombre de bonne prédiction par la syntaxe (frontière syntaxique majeure = frontière de groupe intonatif) en %.*

Tableau VII : *Nombre de bonne prédiction par la syntaxe (absence de frontière syntaxique = absence de frontière de groupe intonatif) en %.*

Tableau VIII : *Moyenne et écart par rapport à la moyenne (z score) pour les différentes longueurs de GI produites par les sujets à l'Expérience I*

Liste des figures

- Figure 1 : *Les constituants prosodiques selon Selkirk (1984)*
- Figure 2 : *Structure prosodique inspirée de Glanzer(1976) et Cruttenden (1986)*
- Figure 3 : *Comparaison des structures prosodiques phonologique et phonétique.*
- Figure 4 : *Oscillogramme, F_0 et durée relative de l'exemple (3), produit par un locuteur natif de l'anglais.*
- Figure 5 : *Oscillogramme, F_0 , durée relative et amplitude de l'exemple (3), produit par un locuteur natif de l'anglais.*
- Figure 6: *Structure prosodique de l'exemple (3).*
- Figure 7 : *Oscillogramme, F_0 et durée relative de l'exemple (4), produit par un locuteur natif du français.*
- Figure 8 : *Structure prosodique de l'exemple (4).*
- Figure 9 : *Courbes intonatives ayant différentes pentes.*
- Figure 10 : *Une frontière de GI marquée par un changement de direction sur la dernière syllabe.*
- Figure 11 : *Deux GI de pentes opposées.*
- Figure 12 : *Retour complet à la fréquence de départ (reset) = deux GI.*
- Figure 13 : *Retour partiel impliquant une seule syllabe = un seul GI*
- Figure 14 : *Retour partiel impliquant plus d'une syllabe = deux GI*
- Figure 15 : *Fenêtre de 20 ms, avancement de 20 ms*
- Figure 16 : *Fenêtre de 25 ms, avancement de 25 ms*
- Figure 17 : *Fenêtre de 25 ms, avancement de 20 ms*
- Figure 18 : *Exemple d'écran d'analyse.*
- Figure 19 : *Pourcentage de production spontanée en fonction du nombre de syllabes dans le groupe intonatif et pourcentage cumulatif.*
- Figure 20 : *Calcul du centre perceptif (P-center)*

Figure 21 : *Exemple de l'écart entre les syllabes selon les centres perceptifs*

Figure 22 : *Exemple de stimulus contenant 9 syllabes avec frontière de GI*

Figure 23 : *Exemple de stimulus contenant 12 syllabes sans frontière de GI*

Figure 24 : *Exemple de reproduction correcte d'un stimulus par un sujet*

Figure 25 : *Exemple de reproduction incorrecte d'un stimulus par un sujet*

Figure 26 : *% cumulatif en fonction du nombre de syllabes dans le GI en parole spontanée et en tâche de rappel*

Liste des sigles et abréviations

F_0 = fréquence fondamentale

syll. = syllabe

GR = groupe rythmique

GI = groupe intonatif.

ST = semi-tons

sec. = seconde

Remerciements

Merci à toute l'équipe du Laboratoire de Sciences Phonétiques pour l'aide apportée à la réalisation de ce projet.

Ces travaux ont été réalisés avec l'aide de fonds alloués par le Conseil de Recherche en Sciences Humaines du Canada par l'entremise d'une bourse d'études supérieures.

1. Introduction : La prosodie dans son ensemble

1.1. Types de prosodie

Le terme prosodie désigne les schèmes suprasegmentaux créés par les variations de trois paramètres acoustiques dans la parole, soit la fréquence fondamentale (F_0), la durée et l'intensité. Les schèmes créés par ces paramètres peuvent être regroupés selon le type d'information qu'ils permettent de véhiculer et l'utilisation qui en est faite : identificatrice, expressive ou structurante (Vaissière, sous presse). La prosodie dite identificatrice transmet de l'information extralinguistique relative aux caractéristiques physiques (homme, femme, enfant) ou psychologiques (lassitude, enthousiasme...) du locuteur et échappe au contrôle de ce dernier. La prosodie expressive porte plutôt de l'information paralinguistique telle que l'affect (joie, tristesse...) et les intentions de communication comme l'emphase et la modalité (énoncé déclaratif, interrogatif, impératif...). Finalement, la prosodie structurante véhicule de l'information linguistique et regroupe, entre autres, des phénomènes tels que l'accent lexical et la structure informationnelle. Une description plus détaillée de ce type de prosodie (structurante) sera donnée dans les pages suivantes, puisque c'est sur cette utilisation de la prosodie que portent les présents travaux.

1.2. Constituants de la prosodie phonologique et phonétique.

Dans le but d'isoler les constituants de la prosodie structurante, il faut d'abord clarifier la définition de certains termes. La terminologie la plus souvent utilisée est celle établie par Selkirk (1984) pour la phonologie de l'anglais. Suivant cette terminologie, les constituants de la prosodie structurante sont organisés à l'intérieur d'une structure hiérarchique représentée dans un schéma arborescent (voir Figure 1, p. 2). Une structure quasi identique a aussi été établie par Beckman et Pierrehumbert (1986), qui ont remplacé le terme *Phonological Phrase* par *Intermediate Phrase*.

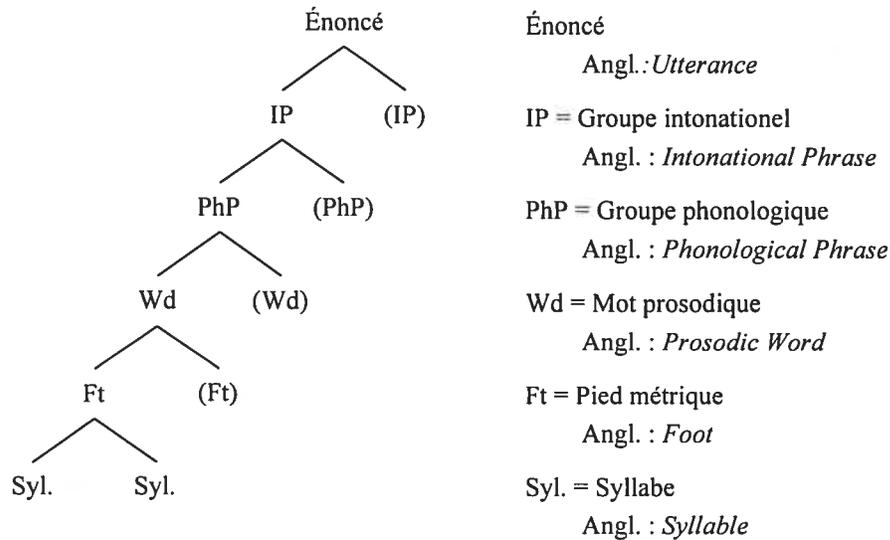


Figure 1 : *Les constituants prosodiques selon Selkirk (1984)*¹

Toutefois, ces constituants prosodiques, issus de théories ou de conceptions phonologiques, ne sont pas définis uniquement en termes acoustiques, mais aussi en fonction de leur contenu. Par exemple, le pied métrique (*Foot*) est considéré comme :

« a suprasyllabic unit, usually smaller in size than the word, that has played a central role in the description of stress patterns in the framework of “metrical phonology” [...] in representing the distinction between stressed and stressless syllables » (Selkirk 1984, p. 31)

Autrement dit, le pied métrique est l’outil qui sert à déterminer l’accentuation primaire (lexicale) d’un « mot ». Selon de telles définitions, les constituants prosodiques ne peuvent être délimités simplement en fonction de ce qui est perçu dans la parole, mais ils nécessitent une interprétation sémantique. De plus, aucun paramètre acoustique ne permet de différencier certains constituants entre eux. C’est le cas du mot prosodique (*Prosodic Word*) et du groupe phonologique (*Phonological Phrase*) qui sont définis respectivement comme : « a roughly word-sized unit of phonological representation » (idem p. 30), et « any level of prosodic constituent structure that may include one or more major category words » (idem p. 29). Un des désavantages de ces définitions

¹ Les traductions proviennent de Millotte (2005)

découle du fait qu'elles sont étroitement liées à la notion de mot qui, en anglais, implique des marques d'accentuation fixe. Toutefois, ces définitions s'appliquent donc difficilement à des langues où le «mot» ne porte aucune marque lexicale, comme le français. Par ailleurs, la seule définition qui réfère à un paramètre acoustique de manière explicite est celle du groupe intonational (*Intonational Phrase*), qui est décrit comme une forme particulière de groupe phonologique caractérisée par une variation de l'intonation (Selkirk 1984, p. 27). Ainsi, vu l'attention portée au contenu linguistique des énoncés, cette terminologie n'est pas applicable à toutes les langues, ce qui limite les généralisations théoriques qui peuvent en être tirées.

On note aussi que cette terminologie a été élaborée en anglais, en fonction de l'anglais, et que plusieurs de ses constituants ne se retrouvent tout simplement pas dans les langues qui ne comportent pas d'accent lexical. Le pied métrique en constitue un bon exemple. Son rôle, tel que défini par Selkirk (1984), est d'attribuer les accents lexicaux. L'existence de ce constituant ne peut donc pas être justifiée par les données du français. Dans cette langue, les mots d'une phrase ne portent pas nécessairement tous un accent. La séquence parlée est plutôt découpée en groupes rythmiques, qui portent à la fois des caractéristiques du «mot prosodique» tel que redéfini par Martin (1987, p. 926) (« [a] minimal segments of prosody that contain one (and only one) stressed syllable ») et du «groupe phonologique», qui peut contenir un ou plusieurs mots lexicaux (Selkirk 1984). L'utilisation de cette terminologie phonologique doit donc être limitée, car elle tend à extrapoler à partir des structures de l'anglais et ne sert pas à désigner les aspects phonétiques de la prosodie structurante en parole spontanée.

Une autre terminologie, inspirée à la fois de Glanzer (1976) et de Cruttenden (1986) et basée sur l'observation de paramètres acoustiques perceptibles, semble davantage applicable au français. Ces deux auteurs postulent les deux seuls niveaux intermédiaires nécessaires à la description du français, soit le groupe rythmique (GR) et le groupe intonatif (GI). De plus, l'emploi de ces constituants prosodiques est facilement généralisable à d'autres langues. Le GR et le GI sont marqués respectivement par un allongement de la dernière syllabe (paramètre « durée ») et par une variation de l'intonation (paramètre « F_0 ») et ne tiennent pas compte du contenu

auquel ils se superposent. Ils sont organisés à l'intérieur d'une structure hiérarchique représentée à la Figure 2 (p. 4).

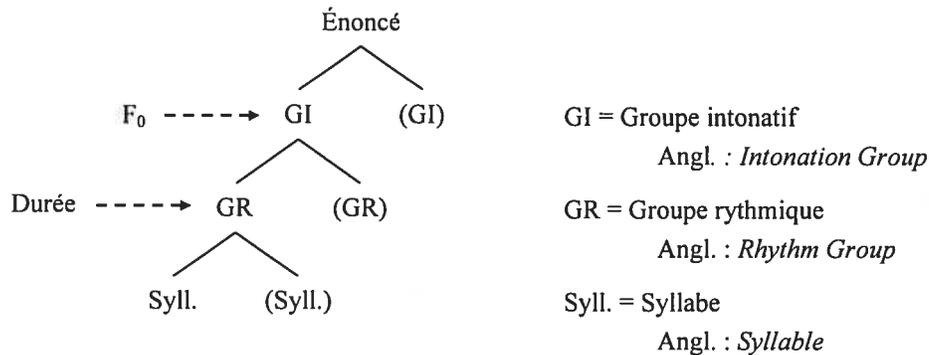


Figure 2 : *Structure prosodique inspirée de Glanzer(1976) et de Cruttenden (1986)*

Il est à noter que cette façon de concevoir la structure prosodique présente l'avantage de convenir aussi aux langues à accent lexical (Nootboom 1997). Bien que les deux structures décrites précédemment présentent des différences importantes, il est possible de faire des parallèles entre certains constituants des deux modèles. Par exemple, les définitions du GI phonétique et du groupe intonational (IP) phonologique en font des constituants semblables. De plus, le GR phonétique a des points en commun à la fois avec le pied métrique, le mot prosodique et le groupe phonologique. La Figure 3, (p. 5) résume les parallèles entre la terminologie phonologique largement inspirée de l'anglais et la terminologie phonétique utilisée dans ce mémoire.

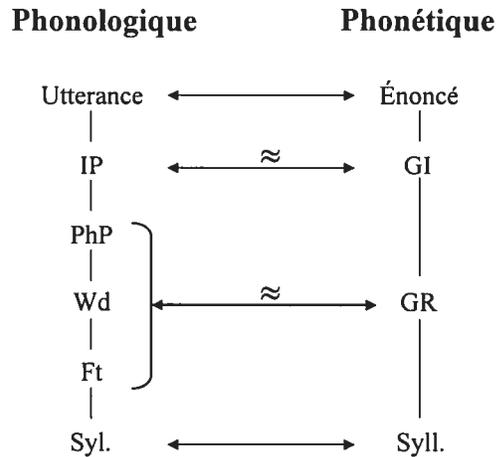


Figure 3 : *Comparaison des structures prosodiques phonologique et phonétique.*

1.3. Avantage du français dans l'étude de la prosodie structurante

Pour étudier les aspects structurants de la prosodie, il est préférable de travailler à partir de langues ne possédant pas de caractéristiques prosodiques codées lexicalement telles que l'accent lexical ou le ton. On s'assure ainsi que les phénomènes prosodiques observés dépendent de fonctions intrinsèques à la prosodie, d'où le critère de « transparence ». Selon ce critère, le français est une langue idéale pour l'étude de la prosodie structurante phonétique car il n'existe à ce jour aucun moyen efficace de prédire l'accentuation et la réalisation des GR et GI de cette langue au-delà de certaines probabilités (Martin 1999). En français ni la phonologie ni la syntaxe ne peuvent prédire efficacement la réalisation des constituants prosodiques de longueur variable, ce qui va à l'encontre des théories phonologiques établies, comme le modèle de Selkirk (1984) présenté précédemment.

1.4. Un problème fondamental : l'incongruité entre la syntaxe et la prosodie

Plusieurs auteurs ont voulu modéliser la prosodie d'énoncés en s'appuyant sur leur structure syntaxique comme facteur d'organisation (voir entre autres Monnin et Grosjean 1993, Hirst 1993, Martin 1991 et Scholes 1971). Toutefois, l'observation de la

parole chez les enfants et la littérature portant sur les effets de longueurs permettent de douter de la pertinence des démarches fondées sur la syntaxe.

1.4.1. Incongruité lors de l'acquisition

D'abord, notons que le nourrisson utilise les structures prosodiques de la langue à laquelle il est exposé dès le stade du babillage (de six à huit mois). Le nouveau-né est donc en mesure de produire les structures prosodiques de sa future langue maternelle bien avant de pouvoir construire ses premières structures syntaxiques (entre quatorze et dix-huit mois chez la majorité des enfants). Comme la prosodie est acquise avant la syntaxe, son apprentissage ne peut pas dépendre de la présence des structures syntaxiques. Il semble donc étrange de supposer que la structure prosodique, acquise très tôt, soit subordonnée à la structure syntaxique, laquelle est acquise beaucoup plus tard (voir entre autres Morley 1957 et Cruttenden 1986).

1.4.2. Effets de longueur

On observe aussi qu'une même structure syntaxique peut mener à la production de structures prosodiques différentes selon la longueur des syntagmes qu'elle contient. Soit les deux phrase suivantes, qui contiennent des sujets de longueur différente :

(1) *Paul part mardi matin*

1 syll

(2) *Alexandrina part mardi matin*

5 syll.

Dans la phrase (1), on a de fortes chances de retrouver un GR contenant à la fois le sujet *Paul* et le verbe *part*. Toutefois, il est quasi impossible que le sujet et le verbe de la phrase (2) soient inclus dans un même GR, et ce, même si ces deux phrases partagent la même structure syntaxique. Il n'y a donc pas de correspondance directe entre structure syntaxique et structure prosodique. Afin de tenir compte de ce phénomène, certains auteurs ont tenté d'ajuster leurs modèles basés sur la syntaxe, sans toutefois réussir à rendre compte de tous les cas de prosodie spontanée (voir entre autres Monnin et Grosjean 1993).

1.5. Fonctions de la prosodie structurante

L'analyse de données statistiques portant sur la longueur moyenne des constituants prosodiques révèle un des rôles possibles de la prosodie structurante de la parole. On observe ainsi que les GR ou les « espaces interaccentuels » qui se retrouvent dans la parole, sont généralement de même longueur que les groupements qui, selon des études en psychologie cognitive, facilitent la mémorisation. Ainsi, Dauer (1983) et Martin (1999) ont démontré que le GR était restreint à 3,55 syllabes en moyenne, et ce, dans plusieurs langues différentes (incluant des langues non indo-européennes). Parallèlement, des psychologues ont démontré que le taux de rappel optimal était obtenu lorsque des stimuli étaient présentés par groupes de trois ou quatre items (Maybery, Permentier et Jones 2002, Reeves, Schauder et Morris 2000, Hitch, Burgess, Towse et Culpin 1996, Frankish 1995,1989). De plus, Boucher (à paraître) a démontré un lien direct entre les effets de groupement sur la mémoire et les GR de la parole par l'entremise d'un test de mémoire sérielle jumelé à une tâche de production quasi spontanée (voir aussi Boucher et Gilbert 2005).

1.6. Conclusion et objectifs

L'effet de la mémoire immédiate sur la longueur des structures prosodiques peut sembler difficile à interpréter dans une perspective linguistique puisque d'éminents auteurs tels que Chomsky ont exclu la mémoire de la compétence langagière, la reléguant au niveau de la performance (Chomsky 1967). Toutefois, l'existence d'un lien entre mémoire et prosodie semble logique lorsqu'on considère qu'apprendre une langue requiert la capacité de reproduire, et donc de mémoriser et de rappeler de nouvelles séries de syllabes structurées prosodiquement. Ainsi, toute restriction sur la mémorisation de la prosodie contribuerait à limiter les structures prosodiques pouvant être acquises à un ensemble limité de patrons plus « faciles » à manipuler sur le plan mnémonique. Les structures plus difficiles ne seraient pas apprises et/ou utilisées et, de ce fait, moins fréquentes en discours spontané.

Dans le but de vérifier si cette prémisse s'applique à des constituants prosodiques autres que le GR, on a effectué une recension de la littérature portant sur la prosodie linguistique phonétique, sur la mémoire sérielle et sur les liens déjà postulés entre ces deux sujets d'étude. La revue de cette littérature a permis d'élaborer une question de recherche et une hypothèse principale sur le lien entre la mémoire et les GI, dans l'optique des effets de longueur décrits ci-dessus. Puisque cette littérature porte majoritairement sur l'anglais, il faut bien saisir les différences entre les phénomènes prosodiques spécifiques à l'anglais et les aspects universaux des structures prosodiques. Ces derniers ont des corrélats acoustiques observables qui permettent de dégager des principes structurants intrinsèques à la prosodie.

2. Recension de la littérature

2.1. Clarification de la terminologie mnémonique

Le terme « mémoire » désigne de façon générale la faculté de conserver et de rappeler des informations avec précision. En psychologie, le concept de mémoire est toutefois divisé en trois sous-systèmes : la mémoire à long terme, la mémoire à court terme et la mémoire de travail.

La mémoire à long terme est associée à la rétention de grandes quantités d'information sur de longues périodes. C'est ce type de mémoire qui est impliqué lorsqu'on apprend le nom des capitales de l'Europe ou le sens du mot « bigoudi » par exemple. À l'opposé, la mémoire à court terme est définie par Baddeley comme étant « the retention of small amounts of information over a brief time interval » (Baddeley 2000, p. 77). L'encodage effectué par ce type de mémoire est plus sujet à l'effacement et disparaîtra rapidement à moins d'être encodé et stocké dans la mémoire à long terme. Parallèlement, le terme « mémoire immédiate » est utilisé pour référer à la faculté de rappeler sur-le-champ des informations qui viennent d'être codées en mémoire (\approx mémoire à « très » court terme). C'est ce type de mémoire qui est impliqué dans l'imitation ou la répétition immédiate. On évalue les performances de la mémoire à court terme d'un sujet en fonction du nombre d'éléments qu'il peut rappeler à partir de listes présentées visuellement ou auditivement. Dans les expériences classiques, ces listes sont constituées de lettres, de chiffres, de syllabes ou de mots. Ces mêmes séries peuvent aussi servir à mesurer la mémoire sérielle si on demande au sujet de rappeler les éléments dans leur ordre de présentation. La mémoire sérielle entre en jeu lorsque l'ordre des éléments à mémoriser est important, comme dans le cas des numéros de téléphone (555-1565 \neq 555-1556) ou lors de l'acquisition du vocabulaire ([pike] \neq [kepi]). Ce type de mémorisation a fait l'objet de plusieurs études en psychologie cognitive, dont les plus pertinentes seront présentées aux sections 2.4.2. et 2.5.4.

Les particularités de la mémoire à court terme et de la mémoire à long terme sont reconnues par la communauté scientifique depuis déjà plusieurs années. Ce n'est cependant pas le cas de la mémoire de travail, qui n'est apparue dans les recherches que dans les années soixante-dix. Son objet n'est donc pas encore clairement établi et

chaque chercheur y va de sa propre définition. Pour cette raison, on évitera d'utiliser le terme « mémoire de travail » dans le présent document et on utilisera plutôt le terme « mémoire immédiate », qui est le terme plus approprié pour désigner les processus mnémoniques étudiés.

2.2. Clarification de la terminologie et des corrélats observables des constituants prosodiques

2.2.1. La superposition des paramètres acoustiques : le cas de l'anglais

On a précédemment introduit les termes de groupe rythmique (GR) et de groupe intonatif (GI) ainsi que leurs corrélats acoustiques : l'allongement et la variation de F_0 . Lors de l'analyse, il faut toutefois aussi tenir compte de la nature hiérarchique de la structure prosodique, qui implique une superposition des paramètres acoustiques associés aux constituants GR et GI tel qu'illustré à la précédente Figure 2 (p. 4). Ainsi, la syllabe finale d'un GI porte à la fois une variation de F_0 et un allongement, car elle se trouve aussi en position finale de GR. De plus, l'interprétation des corrélats acoustiques se révèle encore plus ardue dans les langues à tons ou à accent lexical, où certains paramètres acoustiques codés lexicalement peuvent compromettre l'observation des GR et des GI. Pour illustrer ce phénomène, prenons l'énoncé (3), tel que produit par un locuteur natif de l'anglais (voir Figure 4, p. 9).

(3) *The old contract from that firm requires a signature.*

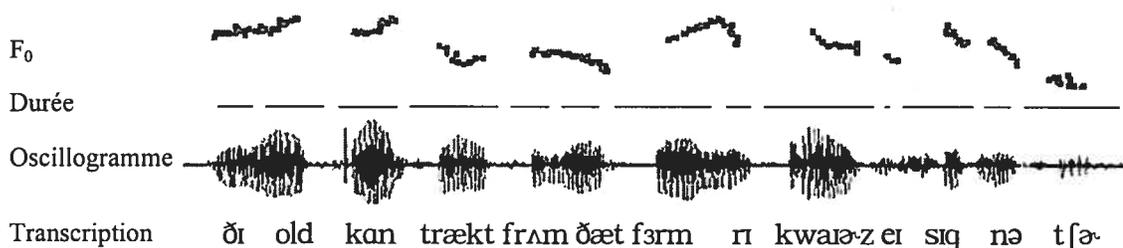


Figure 4 : Oscillogramme, F_0 et durée relative de l'exemple (3), produit par un locuteur natif de l'anglais.

On voit ici une montée abrupte de la F_0 sur la syllabe [fɜrm] suivie d'une baisse sur la syllabe suivante, indiquant une frontière de GI. Un premier GI se termine donc par la syllabe [fɜrm], et un deuxième commence par la syllabe [ɪ]. En déterminant la frontière d'un GI, on délimite automatiquement un GR, vu la position inférieure de ce dernier dans la structure prosodique. Toutefois, on ne peut pas localiser aussi facilement tous les GR, car leur corrélat acoustique, l'allongement, peut être confondu avec les corrélats de l'accent lexical, qui se manifeste aussi, entre autre, par une augmentation de la durée de la syllabe. À preuve, plus de la moitié des syllabes de cet exemple ([old], [kan], [trækt], [fɜrm], [kwæɹ-z], [sɪg] et [tʃə]) semblent « allongées » lorsqu'on compare leur durée à celle des déterminants et des prépositions, qui eux, ne peuvent porter d'accent lexical. Il est donc difficile de retrouver la structure prosodique phonétique des constituants contenant des accents lexicaux sans avoir recours à des analyses qui permettraient de séparer les allongements qui sont dus à ce type d'accent (avec montée d'intensité) de ceux qui signalent la fin d'un GR (sans modification de l'intensité, voir Delattre 1966a). Toutefois, l'ajout d'une analyse de l'amplitude relative des syllabes, telle que présentée à la Figure 5 (p. 10) permet de préciser la nature des accentuations.

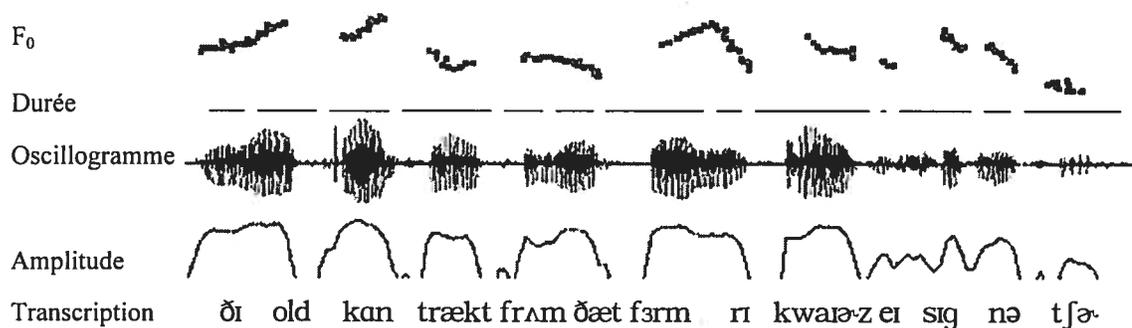


Figure 5 : Oscillogramme, F_0 , durée relative et amplitude de l'exemple (3), produit par un locuteur natif de l'anglais.

Cette troisième analyse permet d'entrevoir que l'allongement des syllabes [trækt] et [tʃə] est causée par la fin d'un GR, car elles ont une moins grande amplitude que la plupart des autres syllabes allongées. Cette analyse doit cependant être

interprétée avec précaution. Par exemple, la syllabe [sɪŋ] ressemble à une fin de GR de par sa durée et son intensité moins élevée, mais cette dernière peut s'expliquer par le degré d'aperture de la voyelle [ɪ] (fermée). De plus, comme le GR ne peut pas séparer les syllables d'un « mot », on ne peut considérer [sɪŋ] comme une frontière de GR. On note cependant que les syllables [trækt], [kwaɪəz] et [tʃə] sont plus longues que les autres syllables allongées. Comme deux de ces syllables sont déjà reconnues comme des fins de GR, on peut étendre cette propriété à [kwaɪəz] et considérer un allongement « plus marqué » comme un corrélat du GR. La nécessité d'utiliser trois analyses pour dégager la structure prosodique d'un énoncé (présentée à la Figure 6, p. 11) démontre qu'il est relativement difficile de reconnaître les constituants prosodiques dans les langues à accent lexical comme l'anglais.

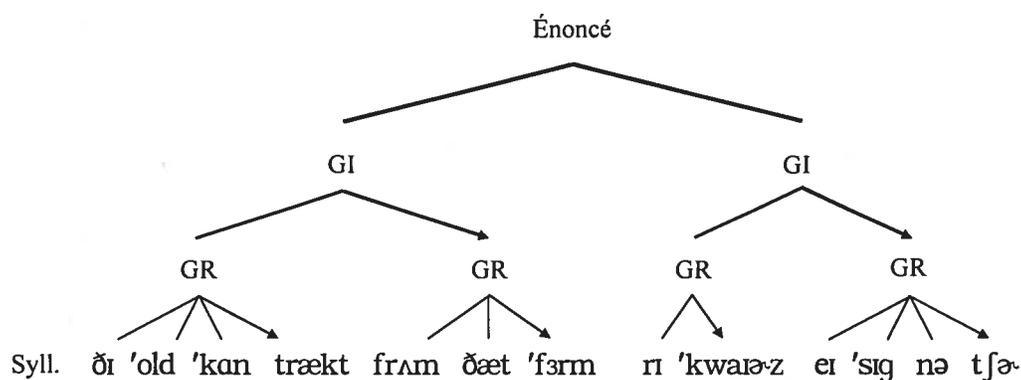


Figure 6 : *Structure prosodique de l'exemple (3).*

2.2.2. La superposition des paramètres acoustiques : le cas (facilitant) du français

Contrairement à l'anglais, le français présente l'avantage de limiter les superpositions de paramètres acoustiques. L'absence d'accent lexical et de ton phonologique dans cette langue permet de dégager la structure prosodique phonétique plus facilement. En effet, comme le GR et le GI ont des corrélats acoustiques différents,

on peut isoler leurs paramètres acoustiques respectifs malgré leur superposition, et ce, sans avoir à tenir compte du paramètre « amplitude ». Pour illustrer cette plus grande simplicité d'analyse, prenons la traduction française de l'exemple utilisé précédemment, soit l'énoncé (4), produit par un locuteur natif du français (voir Figure 7, p. 12).

(4) *Le vieux contrat de cette firme requiert une signature.*

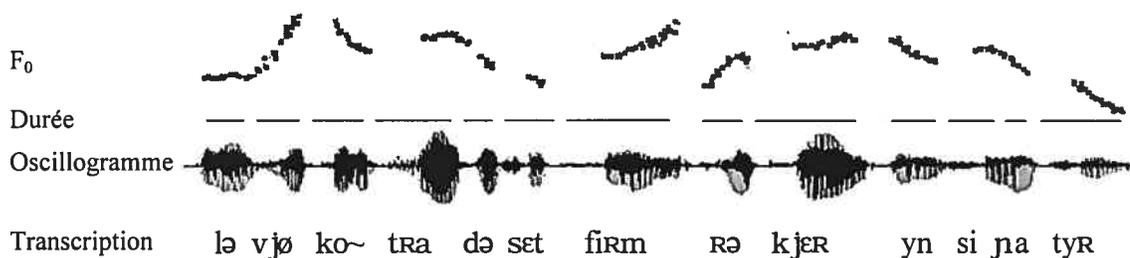


Figure 7 : Oscillogramme, F_0 et durée relative de l'exemple (4), produit par un locuteur natif du français.

On voit ici que les syllabes [tra], [fir̃m], [kjɛr] et [tyr] sont de plus longue durée que les autres. Comme le français n'utilise pas l'accent lexical, ces syllabes représentent nécessairement des fins de GR. On note aussi une montée brusque de F_0 sur la syllabe [fir̃m], indiquant une frontière de GI. Seules ces deux analyses sont nécessaires à la reconstitution de la structure prosodique de l'énoncé, présentée à la Figure 8 (p. 12).

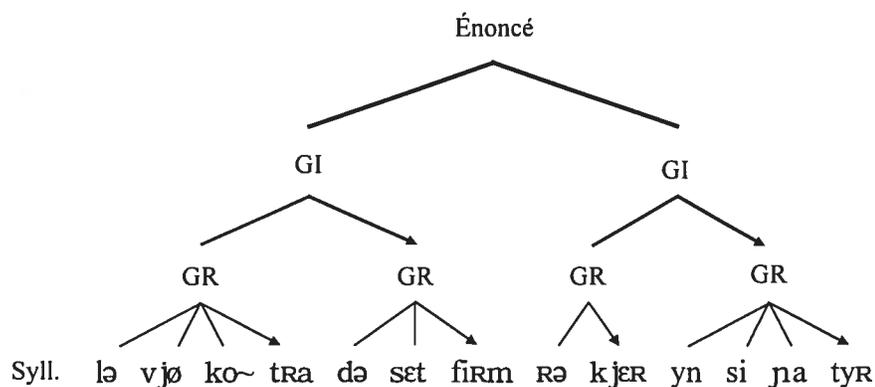


Figure 8 : Structure prosodique de l'exemple (4).

Les constituants prosodiques sont donc plus faciles à localiser en français que dans les langues à accent lexical, malgré la superposition des corrélats acoustiques associés au GI et au GR.

2.3. De la syllabe comme unité de base

Avant tout, il est important d'établir l'unité de base des structures prosodiques. Dans les présents travaux, on considère que l'unité minimale de l'organisation prosodique et de la perception de la parole est la syllabe et non le segment ou le phonème. Cette décision s'appuie sur les résultats de plusieurs études portant sur la production de la parole et sur le biais pouvant être causé par l'écriture alphabétique. La syllabe est d'ailleurs aussi l'unité de base de la hiérarchie prosodique proposée par Selkirk (1984) et d'autres phonologues.

Notons à l'appui que plusieurs auteurs ont tenté de déterminer si la perception des analphabètes était affectée par leur ignorance du système alphabétique. Parmi eux figurent Morais, Cary, Alégria et Bertelson (1979), Byrne et Ledez (1986); Read, Zhang, Nie et Ding (1986) et Ladefoged (1984). Ces auteurs sont tous arrivés à la conclusion que les analphabètes ne sont pas capables de manipuler des unités plus petites que la syllabe. De son côté, Mann (1986) a observé que la majorité des jeunes locuteurs du japonais (langue à syllabaire) n'arrivaient pas à manipuler des phonèmes, ce qui démontre que la perception des segments est reliée à l'écriture alphabétique. De plus, Ehri et Wilce (1980) ont démontré que la connaissance de l'écriture alphabétique influence la perception de suites consonantiques. Ainsi des enfants anglophones ont dit percevoir que le mot *pitch* [pItʃ] contenait un « son » de plus que le mot *rich* [rItʃ]. Seule la présence d'une lettre supplémentaire (*t*) dans l'orthographe de *pitch* permet d'expliquer cette perception, ce qui démontre le biais introduit par l'écriture alphabétique.

Les études sur la production du langage qui remettent en cause la pertinence du phonème sont toutefois moins nombreuses. D'abord, une étude de Boucher (1994) a

démontré, à partir d'un corpus de cinéradiographies, que les erreurs de prononciation résultaient de contraintes empêchant l'atteinte d'une cible articulaire. Le « son » produit par cet « échec » articulaire est perçu comme un autre son, sans toutefois en présenter les caractéristiques. Boucher (soumis) a aussi démontré l'organisation syllabique de la production par une étude électromyographique des muscles d'occlusion et d'aperture de la bouche. Il a ainsi montré que l'ouverture, associée à la voyelle, résulte de la relaxation des muscles de fermeture, et non pas d'une activation distincte des muscles d'aperture. Les mouvements associés aux « consonnes » et aux « voyelles » ne sont donc pas produits en série, mais planifiés simultanément à l'intérieur du cadre de la syllabe.

En somme, un consensus émerge dans la littérature sur le statut de la syllabe comme unité minimale de la parole et de la prosodie. Ce consensus justifie donc le fait de mesurer les « longueurs » de constituants en nombre de syllabes. Ce type de mesure de longueur est aussi utilisé en psycholinguistique et en psychologie dans des études portant sur la mémoire. Le fait de reconnaître la même unité minimale permet alors les recoupements multidisciplinaires entre la linguistique et les sciences cognitives.

2.4. Le groupe rythmique

2.4.1. Le groupe rythmique et ses corrélats phonétiques.

On définit communément le GR comme un ensemble comprenant une syllabe accentuée et toutes les syllabes inaccentuées qui la précèdent (Cruttenden 1986). Plusieurs chercheurs ont tenté d'établir les corrélats acoustiques de l'accentuation. Par exemple, Delattre (1966a) a comparé l'accentuation dans quatre langues (dont le français) en fonction de l'intensité et de la durée relative des syllabes. Il en conclut qu'en français, la durée relative de la syllabe est le seul corrélat acoustique significatif de l'accentuation, les syllabes accentuées étant en moyenne 1,78 fois plus longues que les syllabes inaccentuées. Il a de plus démontré que les augmentations d'intensité étaient réservées uniquement aux accents lexicaux, et ne jouaient donc pas de rôle dans

la prosodie structurante du français. Cet auteur n'a cependant pas examiné le rôle de la F_0 comme marqueur de GR. Selon Cruttenden (1986), les variations de F_0 demeurent réservées à l'accentuation lexicale et n'ont pas de lien avec le GR. D'autre part, Gussenhoven et Rietveld (1992) soulignent que l'importance de l'allongement en finale de GR peut dépendre des autres frontières prosodiques avec lesquelles il coïncide. Ainsi, la dernière syllabe du GR serait toujours allongée, mais cet allongement serait plus important à la frontière d'un autre constituant prosodique, tel que le GI, ce qui concorde avec l'effet de chevauchement discuté précédemment à la section 2.1 (voir aussi Nootboom 1997 et Vaissière sous presse).

Bien que l'existence du GR soit reconnue au sein des langues *syllable-timed* comme le français ou le japonais, sa présence est moins bien acceptée à l'intérieur des langues *stress-timed* telles l'anglais ou l'espagnol. Nootboom (1997) a cependant démontré que l'anglais et le néerlandais présentaient aussi des GR, même si l'allongement final des GR est moins saillant que l'accent lexical, caractérisé aussi par une montée d'intensité (Delattre 1966a). Non seulement le GR est-il présent à l'intérieur des deux grandes familles de langues (*stress-* et *syllable-timed*), mais il y est soumis aux mêmes contraintes de longueur. Ainsi, Dauer (1983) a démontré, en étudiant la production en plusieurs langues de différentes familles linguistiques (thaï, anglais, portugais...) que le GR ou l'« intervalle entre les accents » contenait 3,5 syllabes en moyenne, et 5 syllabes au maximum (significatif). Des résultats similaires ont été observés par Martin (1999) pour le français, à partir d'entrevues radiophoniques dans lesquelles les GR de l'interviewé comportaient en moyenne 3,55 syllabes. Fletcher (1991) a aussi obtenu une majorité de GR contenant entre 3,2 et 4,1 syllabes pour le français (maximum cinq syllabes, minimum deux syllabes).

Cette moyenne située entre trois et quatre syllabes peut être considérée comme représentative de contraintes limitant la production des GR puisque trois auteurs sont arrivés à ces résultats malgré un écart de seize ans entre les expériences de Dauer (1983) et de Martin (1999). Cette contrainte s'applique même si on utilise le modèle des « structures de performance » de Monnin et Grosjean (1993), une théorie basée sur la durée des pauses entre les mots plutôt que sur les variations de paramètres acoustiques.

Ainsi, ces auteurs ont trouvé que le plus petit groupement de leur structure prosodique comportait en moyenne 3,46 syllabes. Cette contrainte se trouve donc légitimée par des données provenant de deux théories différentes, et semble s'appliquer à plusieurs langues.

2.4.2. Sur le rôle structurant du groupe rythmique

Certains linguistes croient que le GR aide l'auditeur à segmenter la parole et qu'il est à l'origine de la perception des « mots » (Cutler 1994). D'autre part, les chercheurs en psychologie connaissent depuis longtemps les effets des GR sur la mémoire sérielle en fonction desquels les sujets participant à des tâches de mémorisation sérielle performant beaucoup mieux lorsque les stimuli sont présentés par groupes de trois ou quatre items. Lorsque le groupement comprend cinq éléments ou plus, les bénéfices mnémoniques disparaissent (voir entre autres Boucher à paraître).

Par exemple, Reeves et coll. (2000) ont démontré que le groupement de séries de chiffres par allongement de la dernière syllabe (avec montée de F_0) améliorait significativement le rappel sériel. Cet effet est moins significatif lorsque la syllabe accentuée se trouve en position initiale plutôt qu'en position finale. Cela démontre que les groupements qu'on retrouve le plus souvent dans la parole (avec allongement final) sont aussi ceux qui permettent une meilleure mémorisation.

Boucher (à paraître) a d'ailleurs établi un lien entre l'utilisation linguistique du GR et son effet sur la mémoire à l'aide d'une tâche de mémoire sérielle et d'une tâche de production dirigée. Il a ainsi démontré que les groupements qui favorisent le plus le rappel sériel correspondent à ceux que les locuteurs ont produits spontanément durant la tâche de production de phrases françaises. Les GR semblent donc remplir une fonction mnémonique pouvant être utile à l'apprentissage et l'interprétation de la parole. Cette étude et celle de Reeves et coll. (2000) suggèrent que ce lien pourrait s'étendre aussi aux structures des GI.

2.5. Le groupe intonatif

2.5.1. Facteurs confondant l'observation des groupes intonatifs

Les variations de F_0 marquant les frontières de GI peuvent être masquées par deux facteurs : la déclinaison et la F_0 intrinsèque des voyelles. La déclinaison consiste en une baisse progressive de la F_0 au cours d'un énoncé. Elle résulte de la diminution de la pression d'air dans les poumons qui entraîne une réduction de la vitesse de vibration des cordes vocales. La déclinaison relève donc de contraintes physiologiques et échappe au contrôle conscient du locuteur.

Par ailleurs, Di Cristo et Hirst (1986) ont démontré, lors d'une étude sur la micromélogie du français, que la F_0 intrinsèque des voyelles est significativement plus élevée lorsqu'elles se trouvent dans une syllabe accentuée dont l'attaque est une occlusive non-voisée. La F_0 de chaque syllabe est donc influencée par la F_0 intrinsèque de la voyelle et par le contexte syllabique, en plus de subir la déclinaison. Il apparaît donc important de ne pas s'attarder à la valeur de fréquence fondamentale de chaque syllabe, mais plutôt d'aborder les variations de la F_0 dans l'ensemble de l'énoncé afin de réduire l'influence des facteurs confondants intrinsèques.

2.5.2. Le groupe intonatif et ses corrélats phonétiques

Bien que Glanzer (1976) ait défini le GI comme « a pattern of pitch, stress, and juncture imposed on a sequence of words » (p. 85), il n'y a pas de consensus quant aux valeurs de F_0 (*pitch*) marquant les frontières du GI. Pour les besoins du présent mémoire, il a donc fallu déterminer des paramètres qui permettraient d'élaborer une définition plus détaillée du GI. Pour ce faire, on s'est fié aux données recueillies lors d'études expérimentales portant sur la perception et la production de la prosodie, en prêtant une attention particulière aux variations de la F_0 .

Dans cet ordre d'idée, Ladd (1983) affirme que les variations momentanées de la F_0 sont perçues par rapport à une valeur de référence (*base line*) établie selon la courbe de déclinaison attendue. Un « retour » à la fréquence de départ (ou *reset*) implique la création d'une nouvelle valeur de référence, et constitue donc un élément important de la

perception des GI. Vaissière (sous presse) abonde dans ce sens en affirmant que plus l'écart entre la F_0 du dernier élément d'un groupe et celle du premier élément du groupe suivant est grand, « plus la sensation de rupture est grande », ce qui favorise la perception d'une frontière de GI. D'autre part, DeZitter (1992) a démontré que la pente de la courbe de F_0 était plus importante que sa hauteur dans la perception de l'intonation. Ainsi, selon ces données, les courbes 1 et 2 de la Figure 9 (p. 18) seraient perçues comme équivalentes malgré qu'elles aient une fréquence de départ différente parce que leur pente est la même. Ces courbes sont toutefois perçues comme différentes des courbes 3 et 4, respectivement moins abrupte et inverse.

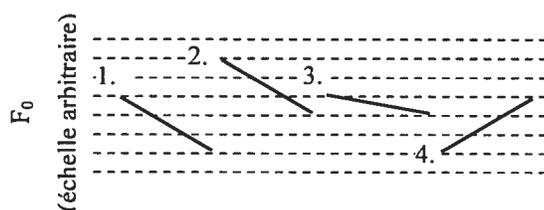


Figure 9 : *Courbes intonatives ayant différentes pentes.*

D'autres auteurs ont tenté de trouver des corrélats articulatoires du GI. Parmi ceux-ci, Fougeron (1999) affirme que les constituants prosodiques sont marqués par un renforcement de l'articulation du premier « segment » visible sur un palatogramme. Byrd et Saltzman (2003), quant à eux, considèrent les variations des niveaux de rigidité et de relaxation comme des paramètres articulatoires influencés par la structure prosodique. Cependant, comme de tels facteurs n'ont pas de corrélats perceptifs confirmés, on n'en tiendra pas compte dans l'établissement des frontières de GI. Par contre, le contenu des GI, c'est-à-dire les unités que ces groupes englobent, constituent un critère accessoire utile pour déterminer les frontières de groupements intonatifs.

2.5.3. Contenu des groupes intonatifs

Les travaux de Fletcher (1991) et de Gussenhoven et Rietveld (1992) s'appuient sur une conception hiérarchique de la prosodie phonétique comme celle décrite à la

section 1.2. En démontrant que la syllabe finale de GI présente un allongement associé à la fin d'un GR, ces auteurs confirment l'existence d'une structure où le GI contient au moins un GR. Cette conception est d'ailleurs partagée par Kenning (1983) selon qui le GI est composé d'un ou de plusieurs GR en français.

En se basant sur cette conception hiérarchique de la prosodie, Sityaev (1999) a présenté des statistiques sur le nombre moyen de GR par GI en anglais, dans des contextes de conversation informelle, de discours informel et de discours formel. Ces résultats indiquaient qu'en contexte non formels, les GI ne contenaient généralement pas plus de deux GR. Les GI contenant plus de deux GR étaient très peu fréquents dans ces styles de parole. D'autre part, le discours formel contenait un bon nombre de GI de trois voire de quatre GR (dans 16,2 % et 10,7 % des cas respectivement). Le discours formel est cependant moins utile à l'étude de la parole car il résulte d'une situation de communication non spontanée. Par conséquent, on ne peut pas considérer ce type de discours comme représentatif de la parole spontanée.

2.5.4. Sur le rôle structurant du groupe intonatif

Il importe de rappeler que les présents travaux portent sur la prosodie structurante et non pas sur la prosodie expressive ou identificatrice. On n'a donc pas tenu compte des études qui ont tenté d'interpréter les courbes intonatives selon leurs fonctions pragmatiques ou émotives (voir entre autres Delattre 1966b, Vaissière 1995, Lacheret-Dujour, Martel et Vivier 1999 et Kenning 1983). Du point de vue de la prosodie structurante, le GI peut être considéré comme un élément facilitant l'interprétation de l'énoncé. Cruttenden (1986), avance d'ailleurs que le GI pourrait servir à mettre la structure syntaxique en valeur.

Notons, par ailleurs, que le GI a aussi fait l'objet d'études en sciences cognitives et en psychologie de la mémoire. Ainsi, Thomas et Hutchens (1990) ont démontré que le rappel sériel de chiffres était facilité par la présence d'une baisse de F_0 sur le dernier item d'une série. Les séquences qui contenaient une variation d'intonation étaient significativement mieux rappelées que leurs équivalents monotones. Frankish (1995) a aussi obtenu de tels résultats qui indiquent que la présence de GI semblables à ceux

qu'on retrouve dans la parole facilitait de façon significative le rappel de séries de chiffres. D'autre part, il a aussi démontré que le fait de superposer la mélodie d'une chanson connue à une série de chiffres à mémoriser ne favorisait pas le rappel sériel. Il conclut donc que la modulation mélodique est inefficace car elle est perçue comme un continuum, et ne crée donc pas d'effet de groupement pouvant faciliter la mémorisation :

«The precise form of the intonation pattern [...] is not crucial. The main requirement is that the auditory cues which mark intonation phrase boundaries should be sufficiently distinctive within the overall pitch contour.» (Frankish 1995, p. 19)

Selon Glanzer (1976), le groupement intonatif facilite l'accès à la mémoire à long terme lors du rappel de séries de mots ayant des liens sémantiques entre eux. Cet effet disparaît si les mots liés sémantiquement se situent dans des GI différents. Par exemple, on obtient de meilleurs résultats lorsque les mots d'un même champ sémantique sont contenus à l'intérieur d'un GI (*chat, chien, rat / table, chaise, pouf / couteau, fourchette, cuillère*) plutôt que dans des GI séparés (*chat, table, couteau / chien, chaise, fourchette / rat, pouf, cuillère*). Le GI semble donc renforcer les liens sémantiques unissant les items qui le composent.

2.6. Les précédents : de la fonction structurante de la prosodie et ses liens avec la mémoire immédiate

Force est de constater qu'un ensemble de travaux portant sur les effets de la prosodie sur la mémoire a largement été mis de côté au cours des dernières années. L'article de Boucher (à paraître) est une des rares études à tenter de démontrer l'existence d'un tel lien depuis plusieurs années. Mais déjà à la fin des années 70, Aldridge (1978) arrivait à la conclusion, après plusieurs expériences sur la perception, que l'interprétation de la parole devait impliquer la mémoire à court terme. « (Aldridge, 1978) has suggested that phonetic analysis may involve short-term memory mechanisms commonly studied in the context of memory research. » (cité dans Aldridge 1981, p. 388)

Dès le début de cette décennie, Leonard (1974) étudiait le lien entre la prosodie et la mémoire verbale. On note que les expériences de cet auteur auraient pu avoir de profondes répercussions sur la conception voulant que la mémoire ne joue aucun rôle dans l'organisation de la compétence langagière (Chomsky 1967). L'apport des résultats de sa recherche n'a toutefois pas été reconnu, et ces travaux n'ont donc pas eu d'influence sur la perspective dominante de l'époque. Dans ses expériences, Leonard (1974) a utilisé une tâche de mémorisation de séries de mots (sans syntaxe) et de phrases (avec syntaxe) produites avec ou sans intonation. Il a ainsi démontré que le rappel des phrases était plus facile que celui des listes de mots, en condition « avec intonation », ce qui semble indiquer que la syntaxe a un effet facilitant la mémorisation immédiate. Toutefois, il a rapporté qu'il n'y avait aucune différence significative entre le rappel des phrases sans intonation et celui des listes de mots. Ainsi, en l'absence d'intonation, l'effet facilitant de la structure syntaxique disparaît. Cela suggère que *la structure prosodique a un effet de plus grande portée sur la mémoire que la structure syntaxique*.

Une expérience du même genre a été menée par Rosner, Grabe, Nicholson, Owen et Keane (2003) dont les stimuli étaient composés de phrases de longueur variable dont on avait manipulé les paramètres prosodiques (avec/sans pause, avec/sans intonation). On voulait ainsi tester l'effet des constituants prosodiques sur le rappel de phrases et la portée de cet effet sur des phrases de longueur variable. Cette expérience aurait pu fournir de précieuses données si on avait contrôlé certaines variables confondantes. Dans le but d'alléger la tâche de rappel, les auteurs ont séparé les stimuli en deux groupes, et chaque sujet n'était alors exposé qu'à la moitié des stimuli. On n'a cependant pas tenu compte du niveau de difficulté des stimuli lors de leur répartition, et on a ainsi créé un groupe significativement plus difficile que l'autre. La mise en commun des données provenant des différents sujets n'était donc plus possible, et les résultats n'ont pu être analysés comme prévu. Il n'y a donc pas de résultats concluants à cette étude bien que son design mérite qu'on y prête attention.

Dommergues et Grosjean (1981) ont adopté une autre approche dans l'étude du lien entre les structures de performance (structures prosodiques) et la mémorisation de phrases. Selon leur hypothèse, le pourcentage d'erreurs dans le rappel des mots d'une

phrase devrait être plus grand en frontière de constituants prosodiques (majeurs ou mineurs) qu'à l'intérieur de ceux-ci. La mesure utilisée était le TEP (*Transitional Error Probability*), soit la probabilité d'erreur transitionnelle calculée selon le nombre de fois où un mot a été mal rappelé, divisé par le nombre total de mots dans la phrase. Ils ont ainsi trouvé que les mots en frontière de constituants prosodiques (selon la « structure de production ») avaient des TEP plus élevés que les mots situés à l'intérieur de ces mêmes groupes *mais seulement lorsque les constituants syntaxiques de la phrase étaient de longueur équivalente*. Certains facteurs, comme l'utilisation du mot comme unité minimale, ont pu influencer les résultats de cette étude mais il ne serait pas pertinent d'entrer ici dans ces détails.

2.7. Conclusions : le rôle structurant des groupes intonatifs et les leurs liens avec les processus de mémorisation.

En résumé, on voit que les notions de GI et de GR présentent des caractéristiques phonétiques observables, mais sujettes à certains facteurs confondants. Ces constituants sont également structurés hiérarchiquement, en ce qu'un GI contient au plus deux GR. Une telle structure implique que des paramètres acoustiques, comme l'allongement marquant la fin d'un GR, peuvent se superposer à des changements de F_0 marquant la fin d'un GI. En ce qui a trait à la délimitation des GI, les travaux recensés précédemment indiquent qu'ils peuvent être marqués par deux critères :

1) Un changement de direction de la pente. Ce critère peut s'appliquer sur la dernière syllabe du groupe pour en marquer la frontière, ou entre deux GI ayant des pentes d'intonation opposées, pour les séparer. Les Figures 10 et 11 (p. 24) illustrent les deux contextes où ce critère s'applique.

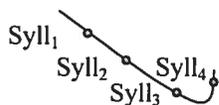


Figure 10 : Une frontière de GI marquée par un changement de direction sur la dernière syllabe.

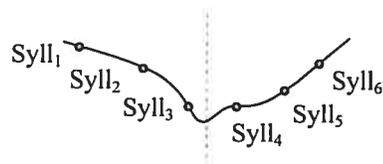


Figure 11 : Deux GI de pentes opposées.

2) Le retour de la F_0 (reset). Ce critère se traduit par une montée soudaine de la F_0 . Cependant, il doit être appliqué en tenant compte de la déclinaison qui limite les modulations de l'intonation (voir section 2.4.1). On considèrera donc qu'il y a retour de F_0 lorsque plus d'une syllabe est produite au-dessus du niveau de F_0 atteint juste avant la variation de fréquence. En fonction de ce critère, les retours de F_0 portant sur une seule syllabe ne seront pas considérés, car ils peuvent provenir de facteurs confondants comme la F_0 intrinsèque des voyelles. Les figures suivantes (12 à 14, p. 23) illustrent l'application de ce critère en fonction des restrictions qu'on vient de décrire.

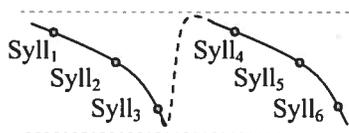


Figure 12 : Retour complet à la fréquence de départ (reset) = deux GI.

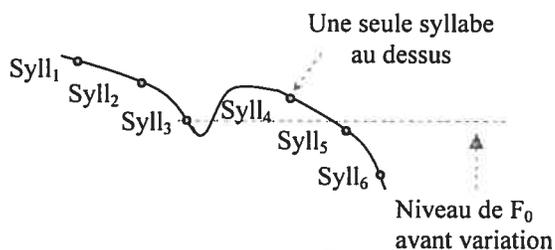


Figure 13 : Retour partiel impliquant une seule syllabe = un seul GI

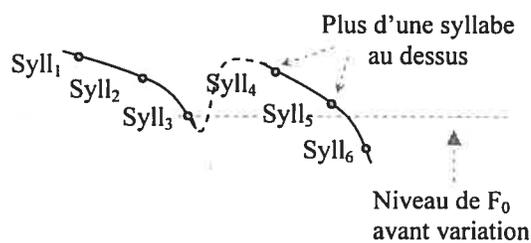


Figure 14 : Retour partiel impliquant plus d'une syllabe = deux GI.

Quant aux aspects structurants des GI, la recension effectuée précédemment suggère que des phénomènes de longueur intrinsèques opèrent indépendamment de la syntaxe et constituent un attribut fondamental de ces schèmes intonatifs. D'autres études avancent que ces effets de longueur dans la parole spontanée peuvent être liés à des processus de mémoire immédiate. L'objectif général du présent travail est de déterminer, dans le cadre d'expériences contrôlées, l'existence d'un lien entre les GI produits en parole spontanée et la mémoire immédiate pour les structures prosodiques. La formulation de cet objectif s'appuie à la fois sur les effets des GI sur la mémoire sérielle présentés dans les pages précédentes et sur le lien déjà démontré entre les GR de la parole et les effets des GR sur le rappel sériel.

3. Hypothèses

Bien que l'existence d'un lien entre le GR et la mémoire ait déjà été démontrée, aucune étude n'a encore démontré qu'un tel lien s'appliquait également au GI. On peut, se demander si les mêmes contraintes de longueur s'appliquent à la fois à la mémoire immédiate et à la production spontanée des GI en français québécois. Plus précisément, on avance l'hypothèse que les restrictions sur la longueur des GI de la parole spontanée correspondent aux contraintes sur la mémoire immédiate pour les GI. La vérification de cette hypothèse se fera en trois étapes, correspondant aux expériences I, II et III ci-dessous.

Premièrement, des statistiques devront être établies sur la longueur moyenne des GI en français québécois. Alors qu'il existe déjà des études statistiques portant sur le GR, la seule étude qui ait pour objet le contenu des GI est celle de Sityaev (1999). Toutefois, en combinant les résultats des travaux sur la longueur moyenne des GR (3,5 syllabes) à ceux de Sityaev (1999) (maximum de deux GR par GI), on peut prévoir que la longueur moyenne des GI se situera entre 3,5 et 7,1 syllabes. L'Expérience I permettra donc à la fois de vérifier cette prédiction et d'établir les statistiques nécessaires à la conception et à l'interprétation des Expériences II et III.

Une fois les statistiques relatives au GI établies, on pourra vérifier si les contraintes de longueur trouvées dans l'Expérience I s'appliquent aussi à la mémoire immédiate pour les GI (Expérience II). Pour ce faire, on a conçu une tâche de mémoire immédiate impliquant des stimuli constitués de structures prosodiques sans contenu linguistique (séquences de la syllabe [pa]) avec F_0 contrôlée. Saito (2001) a légitimé cette façon d'isoler la prosodie en démontrant qu'il y avait une corrélation directe entre la mémoire auditive, qui est utilisée pour le traitement de la parole, et la mémoire pour les rythmes. Dans cette perspective, on prévoit que les contraintes sur la mémoire immédiate pour les GI concorderont avec la limite de longueur des GI en parole spontanée. Autrement dit, les GI fréquents en parole devraient être faciles à mémoriser, tandis que les GI rares ou impossibles devraient être difficiles à rappeler.

Enfin, on vérifiera l'applicabilité de la contrainte sur la longueur des GI comme facteur de prédiction de la prosodie en production dirigée (Expérience III). Dans cette expérience, on comparera l'efficacité des prédictions basées sur cette contrainte à celles qui sont basées sur la syntaxe. Pour ce faire, on utilisera une tâche de production dirigée portant sur des phrases dont la longueur des constituants varie en nombre de syllabes. Si les résultats des deux premières expériences confirment les hypothèses émises, la contrainte de longueur devrait être plus efficace que la syntaxe pour prédire la longueur des GI.

4. Expérience I

Pour tester les hypothèses émises à la section précédente, on doit d'abord déterminer la longueur moyenne et la longueur maximale (en syllabes) des GI produits en parole spontanée. Bien que ces statistiques soient disponibles pour les GR, aucune étude n'a encore établi ces valeurs pour le GI. L'Expérience I a pour but de fournir les statistiques nécessaires à la vérification de l'hypothèse principale. Ces résultats serviront à la fois de référence pour construire les stimuli de l'Expérience II (sur la mémoire) et de point de comparaison pour établir un lien entre la mémoire et la parole spontanée.

Bien qu'aucune étude ne donne de statistiques sur la longueur en syllabes des GI, on peut tout de même prévoir leur longueur moyenne en parole spontanée en se basant sur les résultats d'autres études. En effet, Dauer (1983) et Martin (1999) ont tous deux établi à 3,55 le nombre moyen de syllabes par GR, et ce, dans plusieurs langues. D'autre part, Sityaev (1999) a démontré, en étudiant l'anglais, que la grande majorité des GI ne contenaient pas plus de deux GR en contexte non formel. En combinant les données statistiques sur la longueur moyenne des GR et le nombre moyen de GR par GI, on peut estimer que la majorité des GI contiendront entre 3,55 et 7,1 syllabes. La présente expérience permettra donc de déterminer si les GI en français québécois ont un comportement similaire à ceux observés dans d'autres langues.

4.1. Sujets

Quinze locuteurs natifs du français québécois de 20 à 32 ans (moyenne : 26,7 ans) ont participé à l'expérience. Tous les sujets étaient de sexe masculin, pour faciliter l'extraction de la fréquence fondamentale. Outre la langue maternelle et le sexe, les problèmes d'élocution ou d'audition (présents ou passés) constituaient les seuls critères d'exclusion. Des données concernant la région d'origine et le niveau de scolarité ont aussi été obtenues au moyen d'un questionnaire informel, mais une analyse post-hoc a démontré que ces variables n'avaient aucun effet significatif sur les résultats obtenus.

4.2. Procédure

La tâche des sujets consistait à raconter de façon spontanée une anecdote véridique s'étant produite dans différentes circonstances (à l'école, à l'hôpital, au travail, en voyage et dans un bar). Les thèmes étaient dévoilés un à la fois, et les sujets avaient quelques minutes pour se préparer avant le début de chaque récit, lesquels devaient durer au moins trois minutes. En tout, les sujets ont produit en moyenne vingt minutes de parole spontanée.

La parole des sujets a été enregistrée à l'intérieur d'une cabine à l'épreuve du bruit en utilisant un microphone (Audio-Technica, modèle AT803b) placé à 15 cm et à 45° de la bouche du sujet. Durant la tâche, le sujet était assis à une table, face à un mur, et l'expérimentateur était assis derrière lui. De cette façon, on évitait les effets liés à la communication non verbale entre le sujet et l'expérimentateur. La parole a été enregistrée numériquement à l'aide d'une carte de son externe CSL 4400 (Kay Elemetrics) de 16 bits, à un taux d'échantillonnage de 44,1 kHz. Cet instrument était situé à l'extérieur de la cabine fermée où avait lieu l'enregistrement.

4.3. Analyse

Les données sur la F_0 ont été obtenues à l'aide des « routines de détection de pics » (*peak detection routines*) de CSL 4400 (Kay Elemetrics). Ces routines, qui consistent en des calculs appliqués à l'intérieur d'une fenêtre de prélèvement, qui avance progressivement dans le signal. La durée de la fenêtre, de même que l'intervalle entre ses applications (avancement de la fenêtre), sont contrôlés par l'observateur. Dans le cadre de la présente étude, les paramètres de longueur et d'avancement de la fenêtre de prélèvement ont été ajustés de façon à superposer trois analyses différentes (afin d'éviter les biais pouvant provenir de l'ajustement inadéquat d'un de ces paramètres). Les valeurs utilisées sont les suivantes :

1. Longueur de la fenêtre = 20 ms / avancement de la fenêtre = 20 ms
(Calcul appliqué sur une fenêtre de 20 ms à toutes les 20 ms, voir Figure 15, p. 29)



Figure 15 : Fenêtre de 20 ms, avancement de 20 ms

2. Longueur de la fenêtre = 25 ms / avancement de la fenêtre = 25 ms
(Calcul appliqué sur une fenêtre de 25 ms à toutes les 25 ms, voir Figure 16, p. 29)



Figure 16 : Fenêtre de 25 ms, avancement de 25 ms

3. Longueur de la fenêtre = 25 ms / avancement de la fenêtre = 20 ms
(Calcul appliqué sur une fenêtre de 25 ms toutes les 20 ms, voir Figure 17, p. 29)

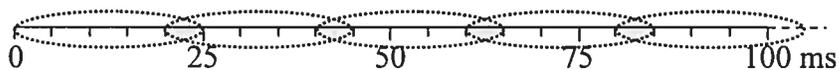


Figure 17 : Fenêtre de 25 ms, avancement de 20 ms

Un chevauchement entre les fenêtres d'analyse permet de compenser la présence dans le signal de coupures causées par les fenêtres de calcul. Les données ainsi obtenues ont été représentées sur une échelle logarithmique, et seuls les résultats qui concordaient dans au moins deux analyses ont été utilisés pour délimiter les GI. De plus, on a superposé le résultat des analyses de F_0 à un spectrogramme à bande large (Blackman, 512 points) pour faciliter le dénombrement des syllabes contenues dans les GI. La Figure 18 (p. 30) illustre cette procédure d'analyse.

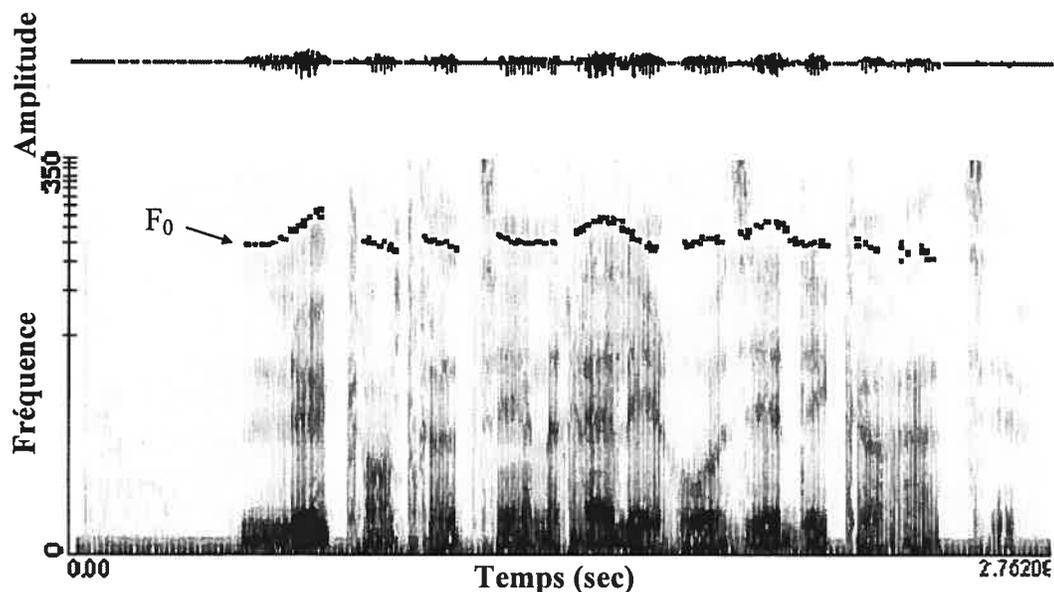


Figure 18 : *Exemple d'écran d'analyse.*

Seuls les énoncés complets, neutres et affirmatifs ont été conservés pour l'analyse. On n'a pas tenu compte des énoncés pour lesquels les mesures de F_0 n'étaient pas claires à cause de certains phénomènes comme un débit rapide ou des suites de consonnes occlusives. On a également éliminé les énoncés dont la F_0 était utilisée à des fins discursives ou pragmatiques (emphase, question, etc.), du fait que cette étude porte sur les aspects structurants de la prosodie. Ces critères de sélection ont été appliqués afin de pouvoir comparer les résultats obtenus à ceux des études recensées précédemment (Dauer 1983, Martin 1999, Sityaev 1999), bien que celles-ci ne fassent pas explicitement du traitement réservé aux énoncés incomplets ou non standards. Cependant, en examinant les méthodes utilisées dans ces travaux, on déduit que ces types d'énoncés devaient y être beaucoup moins fréquents que dans la présente étude. En effet, les tâches privilégiées dans ces études sont des tâches de lecture (Dauer 1983 et Sityaev 1999), des entrevues radiophoniques formelles (Martin 1999) ou la récitation de textes appris par cœur (Sityaev 1999). Ces tâches ne font pas appel à la parole spontanée et sont donc moins propices à la production de phrases naturelles.

Par ailleurs, on a établi les frontières de GI en fonction des écarts de fréquences et des changements de direction des pentes telles que décrit dans l'introduction à la section 2.6. On a compté les syllabes des GI en écoutant des portions du signal et en utilisant la procédure d'analyse illustrée à la Figure 18, p. 30.

4.4. Résultats

Les sujets ayant produit en moyenne vingt minutes de parole spontanée libre, un total de 1443 GI ont pu être analysés. De plus, on retrouve une distribution normale des résultats pour chaque longueur de GI à travers les sujets ($z < 2,579$), ce qui a permis de regrouper toutes les données en une seule analyse. (Les résultats détaillés par chaque sujet sont présentés en Annexe.) En moyenne, les GI analysés contenaient 5,27 syllabes, pour un maximum de huit syllabes (limite statistique atteignant le 95^e percentile (94,89), voir Figure 19, p. 31). Le taux de fiabilité des mesures a été établi à 93,1 %, après réanalyse de 10 % des GI (135/145). De plus, dans neuf cas sur dix, l'écart entre les mesures divergentes était inférieur à deux syllabes.

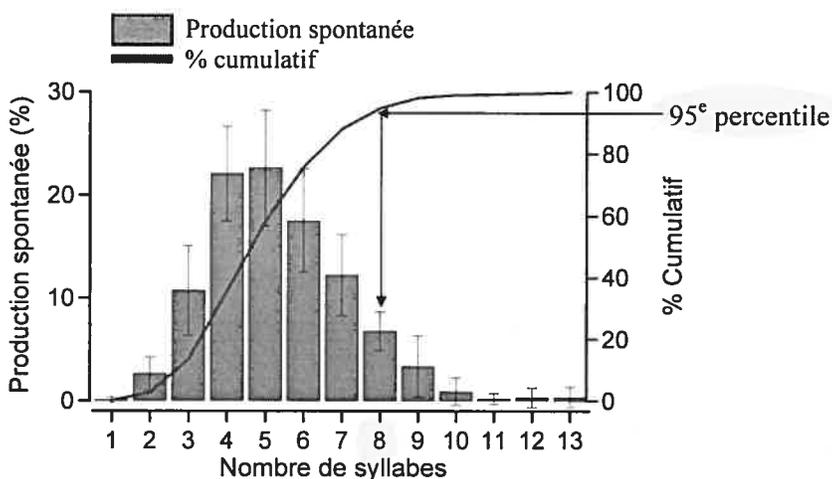


Figure 19 : Pourcentage de production spontanée en fonction du nombre de syllabes dans le groupe intonatif et pourcentage cumulé.

4.5. Discussion

Cette moyenne globale de 5,27 syllabes par GI confirme l'hypothèse secondaire, selon laquelle la longueur moyenne des GI du français québécois devrait se situer entre 3,55 et 7,1 syllabes. En fait, 85,34 % des GI répertoriés contiennent entre trois et sept syllabes, ce qui signifie qu'une forte proportion des GI se situe à l'intérieur des limites de longueur extrapolées à partir des données de Dauer (1983), Martin (1999) et Sityaev (1999). Ces résultats démontrent donc qu'il y a une contrainte observable sur la longueur des GI en nombre de syllabes.

Rappelons toutefois que l'objectif principal de l'Expérience I était d'établir les statistiques nécessaires pour la conception et l'interprétation des expériences suivantes. On a donc utilisé ces statistiques pour formuler l'Expérience II, qui porte sur les contraintes mnémotiques universelles sur la longueur des GI, et l'Expérience III, qui compare l'efficacité des prédictions basées sur cette contrainte à celles basées sur la syntaxe.

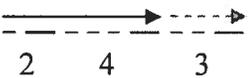
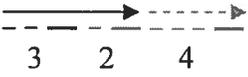
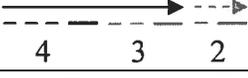
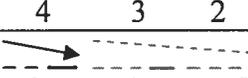
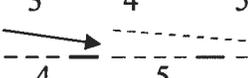
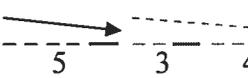
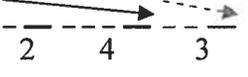
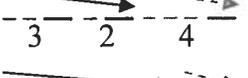
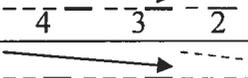
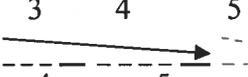
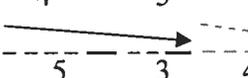
5. Expérience II

Lors de l'Expérience I, on a établi que les GI en parole spontanée ne dépassent généralement pas huit syllabes, réparties dans au plus deux GR (Sityaev 1999). L'objectif de l'Expérience II sera donc de déterminer si cette même contrainte de longueur s'applique aussi au rappel de GI constitués de syllabes répétées ([pa]) groupées rythmiquement.

En soi, le rappel du nombre de syllabes répétées contenues dans des schèmes intonatifs est une tâche relativement facile. On peut penser, par exemple, à la facilité avec laquelle on peut se rappeler la mélodie d'une chanson en langue étrangère, sans toutefois pouvoir se souvenir de son contenu segmental. Cependant, lorsque la longueur des groupements rythmiques et intonatifs dépasse une certaine limite, des erreurs surviennent dans le rappel du nombre exact de syllabes qu'ils contiennent.

Dans cette optique, la tâche de la présente expérience a été conçue afin 1) d'isoler les effets spécifiques des groupements intonatifs sur le rappel du nombre exact de syllabes entendues et 2) de déterminer si les limites s'appliquant au rappel des GI correspondent à la longueur maximale des GI en parole spontanée (Expérience I). Pour répondre à ces deux objectifs, la tâche de rappel impliquait un ensemble de stimuli, c'est-à-dire des séries de syllabes élaborées avec des GI contenant un ou deux GR, pour refléter les structures de la parole spontanée (Sityaev 1999). De plus, il fallait que les stimuli contiennent des GR de longueur variable répartis de façon à éviter les effets de symétrie (jamais deux GR consécutifs contenant le même nombre de syllabes), qui pourraient biaiser le rappel. Dans ces conditions, il fallait tenir compte de l'effet possible des GR sur le rappel des GI. Un ensemble de stimuli a donc été élaboré pour répondre à ces objectifs en exploitant un design particulier. Pour saisir ce design, le Tableau I (p. 34) présente les séries qui ont été utilisées dans la tâche de rappel.

Tableau I : *Stimuli utilisés dans l'Expérience II*. Les flèches représentent de façon schématique les changements d'intonation dans les séquences.

#	Séries élaborées	Condition d'intonation	Possible en parole
1		Neutre	
2		Neutre	
3		Neutre	
4		Avec intonation	✓
5		Avec intonation	✓
6		Avec intonation	✓
7		Avec intonation	✓
8		Avec intonation	✓
9		Avec intonation	✓
10		Avec intonation	✓
11		Avec intonation	✓
12		Avec intonation	✓
13		Avec intonation	✓
14		Avec intonation	
15		Avec intonation	✓

Notons que les séries contiennent soit neuf ou douze syllabes et que toutes sont composées de trois GR de longueur variable (deux à cinq syllabes), en évitant les cas de symétrie. Par ailleurs, certaines séquences contiennent des frontières intonatives qui permettent de délimiter deux GI, alors que d'autres séquences ne contiennent aucune frontière de GI. La tâche était conçue afin de mesurer le rappel exact du nombre de syllabes contenues dans les GI en position initiale, car les GI en position finale peuvent être influencés par la charge mnémonique du GI précédent. Ainsi, un groupe court, qui aurait facilement été rappelé dans d'autres circonstances, pourrait être plus difficile à rappeler s'il est précédé d'un GI long, qui monopolise la mémoire. Des comparaisons de sous-ensembles de stimuli répondent aux deux problèmes principaux et à la nécessité d'examiner les effets possibles des GR sur le rappel des GI.

En particulier, pour isoler l'effet spécifique des frontières intonatives sur le rappel du nombre de syllabes entendues, on a utilisé des séquences de neuf syllabes avec ou sans frontière intonative, mais contenant des GR de même longueur (soit les séquences 1, 2 et 3 comparées aux séquences 10, 11 et 12 du Tableau I, p. 34). On prévoyait que la présence d'une frontière intonative influencerait le rappel du nombre de syllabes contenues à l'intérieur du GI couvrant les deux GR initiaux.

Pour déterminer si le rappel des GI est soumis à la même contrainte de longueur que la production des GI en parole spontanée (en général huit syllabes en deux GR), le rappel des séquences 10 et 11 (de neuf syllabes) a été comparé à celui des séquences 14 et 15 (de douze syllabes). Ces séquences contiennent toutes des GI de longueur différente. D'une part, on a des GI qui sont possibles en parole spontanée (5 ou 6 syllabes), d'autre part, des GI dont la longueur rejoint ou dépasse la limite des GI possibles en parole spontanée (huit ou neuf syllabes). On prévoyait une différence significative entre les performances de rappel des GI en fonction de la longueur de ces derniers et une nette détérioration du rappel pour les séquences dépassant la limite de longueur des GI en parole.

Enfin, la présence de GR de longueur variable dans les GI de la tâche précédente représente un facteur pouvant influencer le rappel. Afin de cerner l'effet de ce facteur, on a utilisé une analyse particulière impliquant un test de Friedman et portant sur

l'ensemble des contextes avec intonation présentés dans le Tableau I (p. 34, séquences 4 à 15). Ce test non paramétrique assigne un « rang » aux résultats obtenus pour chaque longueur de GI initiaux (rappel exact) et permet de comparer leur importance relative. On prévoyait que si les GR avaient une influence secondaire, le rappel s'échelonnerait en fonction de la longueur des GI, et non en fonction de la longueur des GR.

5.1. Stimuli

Les stimuli du Tableau I (p. 34) ont été construits en suivant une technique de resynthèse effectuée au moyen d'un synthétiseur de type Klatt (*ASL 5104*, Kay Elemetrics). Cette méthode préserve l'aspect « naturel » des sons de la parole, tout en permettant la manipulation des paramètres d'intonation et de durée nécessaires à la création de schèmes prosodiques. La resynthèse implique qu'on enregistre d'abord des syllabes, qui seront par la suite modifiées. Dans le cas présent, un locuteur masculin a produit des syllabes [pa], qui ont été numérisées. Les séquences de neuf et de douze syllabes du Tableau I (p. 34) ont été élaborées en assemblant des duplicatas d'une syllabe [pa] sélectionnée et en y apportant des changements de durée et de F_0 au moyen des fonctions du synthétiseur.

Pour créer les changements de durée nécessaires à la perception des GR, on a « allongé » un duplicata de la syllabe d'origine (inaccentuée, 0,192 seconde) par resynthèse jusqu'à l'obtention d'une syllabe de 0,315 seconde. Le ratio de durée entre ces deux syllabes (1,7/1) reflète le ratio d'accentuation du français (Fant, Kruckenberg, et Nord, 1991, Delatre 1966a). À partir de ces deux syllabes, on a assemblé des GR de longueur différente à l'aide des fonctions d'édition de *CSL 4400* (Kay Elemetrics).

Afin de maintenir une isochronie perceptive dans la séquence, on a ajusté la durée des espaces intersyllabiques en fonction du centre perceptif (*P-center*) des syllabes (Marcus 1981). Selon Marcus (1981), la perception de la rythmicité ou de l'aspect isochrone d'une séquence se fait selon le « centre perceptif » des syllabes, et non pas en fonction des durées absolues entre les syllabes. L'emplacement du centre perceptif se

calculé par la formule suivante : $0,65a + 0,25b + c$, où a représente la durée de l'attaque, b , la durée de la rime, et c , une constante dont la valeur a été maintenue à 0 pour la présente étude. La Figure 20 (p. 37) montre comment on a déterminé l'emplacement du centre perceptif de la syllabe accentuée.

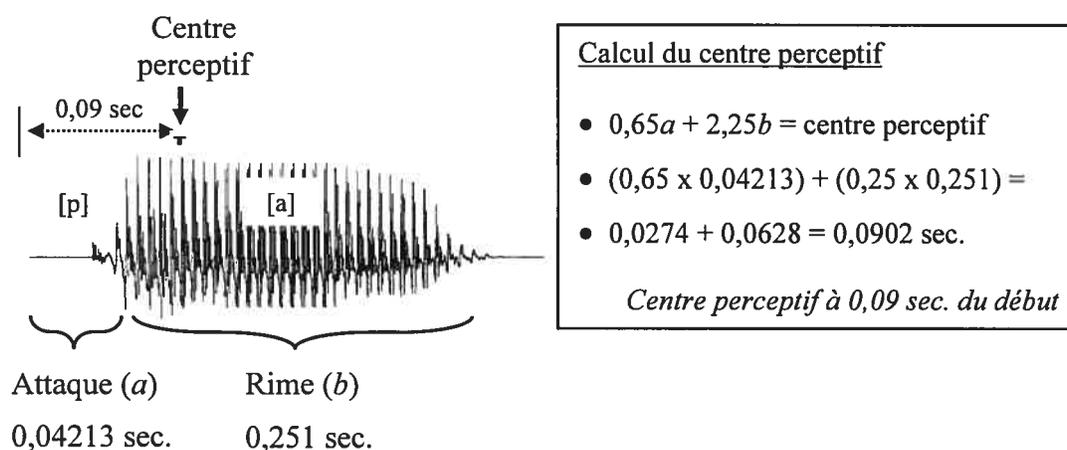


Figure 20 : *Calcul du centre perceptif (P-center)*

L'intervalle entre les centres perceptifs a été fixé à 280 ms à l'intérieur d'un GR, et à 450 ms entre les GR, ce qui donne des séquences où la rythmicité d'accent était clairement perçue. La Figure 21 (p. 37) donne un exemple de ces ajustements en fonction des centres perceptifs.

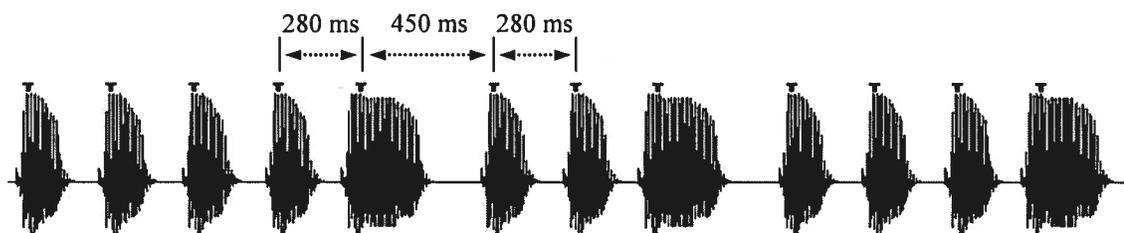


Figure 21 : *Exemple de l'écart entre les syllabes selon les centres perceptifs.* Cette séquence représente des GR de 5, 3 et 4 syllabes avec accentuation en finale de groupe.

Les séquences ainsi construites duraient trois secondes, si constituées de neuf syllabes, ou 3,84 secondes, si constituées de douze syllabes. Leur durée se situait donc à l'intérieur des limites de la mémoire immédiate pour l'intonation (Cole 1973, Springer 1973, Wolf 1977 cités dans Cooper et Sorensen 1981).

Une fois toutes les séquences montées rythmiquement, leur F_0 a été modifiée à l'aide du synthétiseur susmentionné afin de créer les GI. Les frontières de GI ont été marquées par un écart de F_0 (*reset*) correspondant à onze semi-tons (ST) suivant une pente simple descendante. On a utilisé la forme la plus simple de pente d'intonation pour assurer un maximum de saillance à la frontière de GI, tout en respectant les capacités optimales du synthétiseur. Une synthèse sur une gamme plus étendue de fréquences risquait de créer des signaux moins proches de la parole humaine. On a par la suite ajouté un signal avertisseur consistant en un ton pur de 750 Hz, d'une durée de 300 ms, à 1,5 seconde du début du stimulus à l'aide du logiciel *CSL 4400*.

La dernière étape du montage des stimuli synthétisés consistait à diminuer le taux d'échantillonnage afin d'éliminer de possibles artefacts dus à l'édition des signaux. La baisse du taux d'échantillonnage restreint la gamme de fréquences présentes dans un enregistrement en éliminant les hautes fréquences, plus susceptibles de contenir de tels artefacts. Les Figures 22 et 23 (p. 38) donnent respectivement des exemples de stimuli dans leur version définitive, avec et sans frontière de GI (intonation neutre)

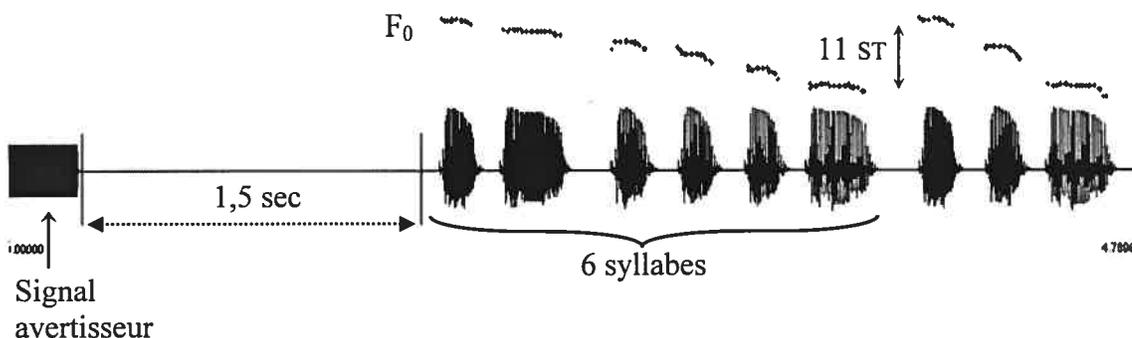


Figure 22 : Exemple de stimulus contenant 9 syllabes avec frontière de GI.

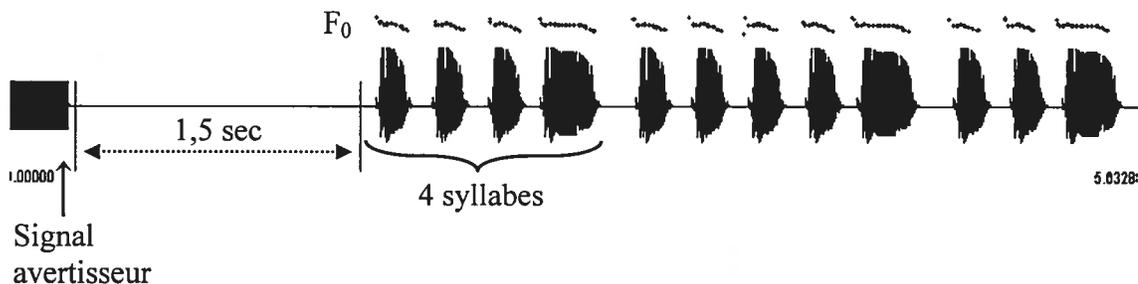


Figure 23 : *Exemple de stimulus contenant 12 syllabes sans frontière de GI*

5.2. Sujets

Quarante-deux locuteurs natifs du français, dix-neuf femmes et vingt-trois hommes, de 19 à 36 ans (moyenne : 24,9 ans) ont été recrutés pour participer à l'expérience. Ces sujets n'ont rapporté aucun trouble d'audition ou de production du langage (présent ou passé). De l'information sur le lieu d'origine et la formation en musique a été recueillie par questionnaire informel. Tous les sujets ont été recrutés au sein de la population étudiante de l'Université de Montréal, et cinq dollars leur ont été offerts en échange de leur participation.

5.3. Procédure

La présentation des stimuli s'est faite par l'entremise d'une présentation *Power Point* (Microsoft), où étaient aussi insérés des séquences servant à une autre expérience. Chacune des séquences était entendue deux fois, à différents moments de la présentation, exception faite des stimuli sans frontière de GI (neutres), qui n'étaient entendus qu'une seule fois. Chaque série de syllabes était précédée d'un avertisseur visuel (*) et d'un signal sonore. De cette façon, chaque sujet était exposé à deux GI de 2, 6, 8 et 9 syllabes et quatre GI de 3, 4, 5 et 7 syllabes, de même qu'à trois séquences neutres, pour un total de vingt-sept réponses analysables par sujet. L'ordre de présentation était aléatoire, suivant une fonction générée par le logiciel *Excel*

(Microsoft). Cet ordre a été ajusté afin d'éviter les suites de stimuli commençant par des GR de même longueur ou des GI contenant le même nombre de GR. Les séquences sans frontière de GI ont été présentées en début de tâche, afin de prévenir l'influence de groupements pouvant être induits par des rythmes perçus précédemment. La durée totale de la tâche ne dépassait pas vingt minutes. La présentation sonore des stimuli s'est faite à l'intérieur d'une cabine à l'épreuve du bruit par l'intermédiaire d'un haut-parleur Studio 20 (Paradigm Electronics) placé à 70 cm, en face du sujet (± 10 cm selon la posture du sujet). L'amplification des stimuli a été réglée de manière à obtenir un maximum de 74 dBA aux oreilles du sujet.

Des instructions écrites étaient données sur moniteur LCD, et la présentation des stimuli ne débutait que lorsque le sujet avait terminé la lecture des directives à voix haute. Les directives se lisaient comme suit :

- Vous allez entendre différentes séries de syllabes.
- Votre tâche consiste à reproduire ces séries de la façon la plus exacte possible.
- Chaque série sera précédée d'un * et d'un signal sonore.
- Trois exemples suivent afin de vous pratiquer...

Notons que les directives étaient volontairement imprécises, afin de ne pas diriger l'attention des sujets vers un aspect prosodique en particulier. Trois exemples servant de pratique ont d'abord été présentés. De plus, l'expérimentateur contrôlait le rythme de présentation des stimuli à partir de l'extérieur de la cabine, tout en maintenant un contact audio avec le sujet. Ce faisant, on s'assurait d'un débit de présentation constant, favorisant l'attention du sujet tout en s'ajustant à sa vitesse de réponse.

Les réponses des sujets étaient captées à l'aide d'un micro-casque omnidirectionnel AKG C477, ce qui permet de capter à la fois le stimulus et la réponse du sujet, ce qui facilite l'étape d'analyse. L'enregistrement se faisait directement sur support numérique, à l'aide d'une carte externe CSL 4400 de Kay Elemetrics (16 bits, 44,1 kHz).

5.4. Analyse

La méthode d'analyse était la même que celle de l'Expérience I. Étant donné que les stimuli ont été enregistrés avec les réponses du sujet, l'expérimentateur avait toujours le stimulus en regard lors de l'analyse et pouvait ainsi évaluer l'exactitude de la réponse, comme l'illustre les Figures 24 (p. 41) et 25 (p. 42). Bien que l'objet principal de cette expérience soit la mémoire immédiate pour les GI, on a aussi tenu compte du rappel des GR pour déterminer si un stimulus avait été rappelé correctement, et ce, dans le but de pouvoir comparer les résultats obtenus à ceux de parole spontanée. Enfin, l'analyse statistique des résultats obtenus a été faite à l'aide du logiciel *SPSS* (version 11.0).

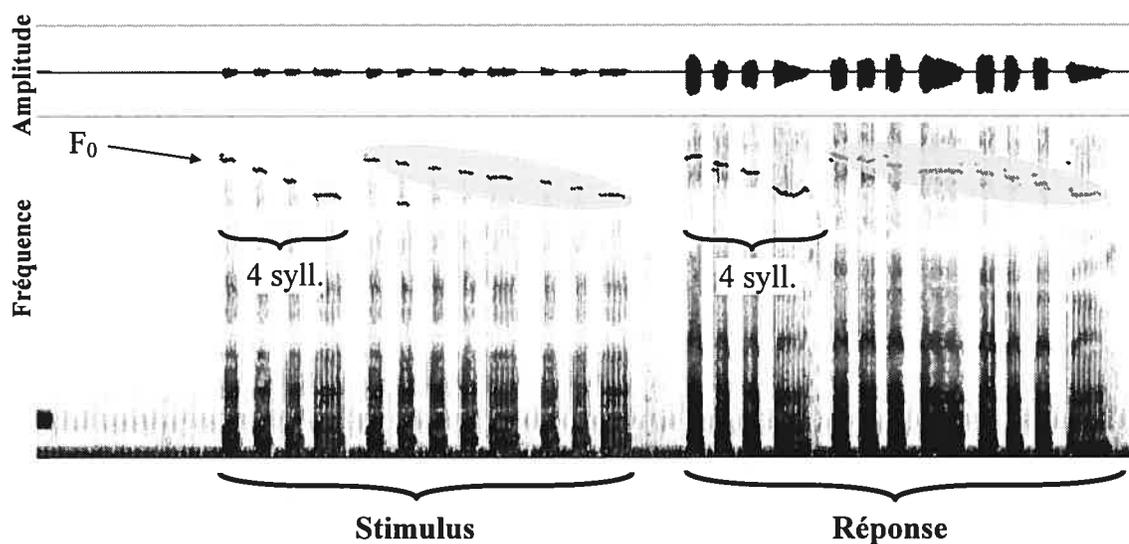


Figure 24 : *Exemple de reproduction correcte d'un stimulus par un sujet.* Notons que la réponse comprend un retour de la F_0 après la 4^e syllabe, ce qui indique un rappel exact du GI dans ce cas.

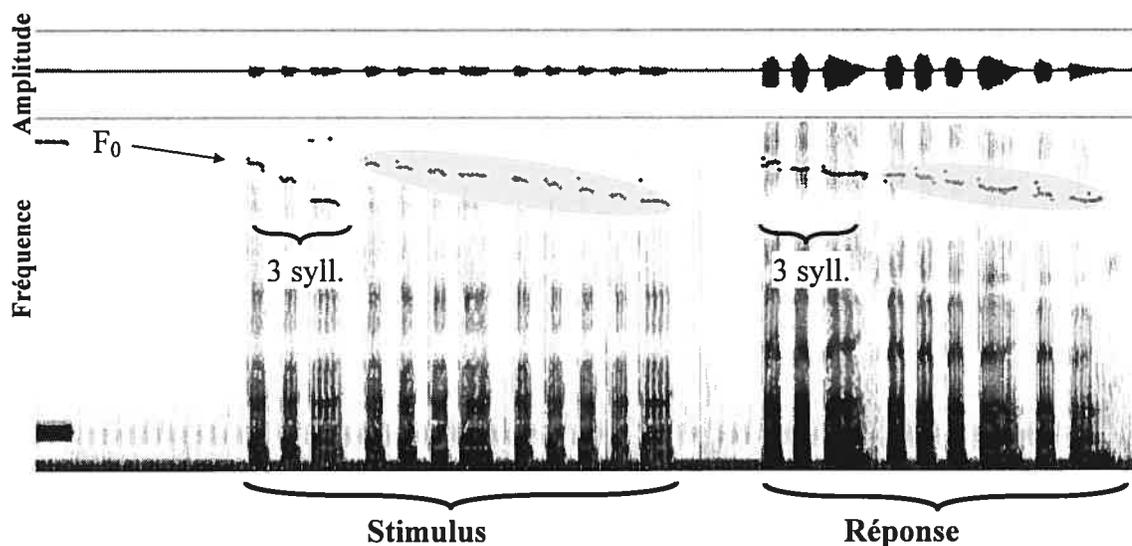


Figure 25 : *Exemple de reproduction incorrecte d'un stimulus par un sujet.* Ici, la réponse du sujet ne contient pas de retour de F_0 indiquant qu'il n'y a pas eu de rappel exact du GI.

5.5. Résultats

On a établi la fiabilité des données en procédant à une réanalyse acoustique de 15 % des enregistrements et le taux d'accord test-retest était de 94,6 % (105/111). Par ailleurs, un examen des données brutes a révélé qu'on ne pouvait assumer la distribution normale des résultats par facteurs, de sorte que des tests statistiques non paramétriques de Wilcoxon ont été appliqués pour la vérification des différentes hypothèses. On a également constaté qu'un sous-groupe de sujets musiciens présentait des performances différentes par rapport aux sujets nonmusiciens. En fait, une comparaison des scores de ces deux groupes sur les huit longueurs de GI a indiqué des différences significatives ($t = 3,863$, $n = 8$, $p = 0,006$) On a donc exclu ces sujets de l'analyse. Les résultats ci-dessous réfèrent donc à trente et un individus, dix-huit hommes et treize femmes, âgés entre 19 et 36 ans (moyenne : 24,5 ans) et dont les scores aux tâches de rappel sont répartis selon une distribution normale ($z < 2,08$).

Pour ce qui est de la première prédiction sur l'effet de la présence d'une frontière intonative quant au rappel du nombre de syllabes des GI, une comparaison des résultats

faite par un test de Wilcoxon a révélé une différence significative entre les deux conditions d'intonation ($z = -3,801$, $n = 31$, $p = 0,000$). Notons que, selon les données statistiques, les GI ne facilitent pas le rappel du nombre de syllabes, mais imposent, au contraire, une charge mnémonique additionnelle.

En ce qui a trait à la deuxième prédiction, voulant que la longueur du GI ait un effet sur le rappel, un second test de Wilcoxon a permis de démontrer une différence significative entre le rappel des groupes probables en parole (5 et 6 syllabes) et le rappel des GI rares (8 et 9 syllabes; $z = -3,828$, $n = 31$, $p = 0,000$).

Enfin, pour évaluer l'influence possible des GR de longueur variable sur le rappel des GI, on a appliqué un test non paramétrique de Friedman à l'ensemble des GI présentés au Tableau I (p. 34). Cette analyse a révélé que les scores de rappel correct du nombre de syllabes dans les GI s'échelonnaient de façon significative selon la longueur des GI et non selon la longueur des GR. Le Tableau II (p. 44) résume ces résultats. Les statistiques de ce tableau permettent de voir la diminution de performance (représentée par le rang) en fonction de la longueur du GI. On remarque, entre autres, que même si un GI comprend des GR courts, son rappel est fonction du nombre total de syllabes qu'il contient (voir aussi les statistiques descriptives du Tableau III, p. 44).

Tableau II : *Test non paramétrique de Friedman, comparant l'ensemble des GI en position initiale de séquence.*

	Description	Illustration du GI	Rang moyen
1 groupe rythmique	2 syllabes		7,74
	3 syllabes		6,85
	4 syllabes		6,26
	5 syllabes		5,19
	5 syllabes		4,71
2 groupes rythmiques	6 syllabes		4,60
	7 syllabes		3,58
	8 syllabes		3,24
	9 syllabes		2,82
<i>n</i> = 31		<i>Sig. Asymp.</i> = 0,000	

Tableau III : *Statistiques descriptives en fonction de la longueur du GI en position initiale.*

longueur du GI (en syllabes)	% du rappel correct	% cumulatif du total
2	96,78	24,84
3	76,61	44,51
4	70,97	62,73
5	46,77	74,74
6	40,32	85,09
7	24,19	91,3
8	19,35	96,27
9	14,52	100

5.6. Discussion

L'expérience ci-dessus visait principalement à déterminer si la contrainte sur la longueur des GI en parole pouvait se refléter dans une limite sur le rappel des schèmes prosodiques. Les résultats obtenus indiquent que le rappel des GI dont la longueur excède celle des GI qu'on retrouve dans la parole (huit ou neuf syllabes) est significativement inférieur au rappel de GI en deçà de la longueur limite trouvée à l'Expérience I (cinq ou six syllabes). Notons que la distribution des scores des sujets ne présente pas de saut marqué pour les GI de huit syllabes (qui suggérerait l'effet d'une limite), mais plutôt une baisse progressive de performance au point où le rappel correct tombe à moins de 15 % pour les GI de neuf syllabes, qui sont quasi inexistantes en parole.

Les résultats démontrent par ailleurs que la présence de frontières intonatives influence le rappel de façon significative, et qu'il est généralement plus difficile de rappeler une séquence *avec* un schème intonatif qu'une séquence *sans* intonation (neutre). Autrement dit, l'intonation ne facilite pas le rappel, mais impose une charge mnémonique additionnelle. En effet, 22 des 31 sujets ont moins bien rappelé les séquences avec frontière de GI que les séquences équivalentes sans frontières.

L'existence d'un effet de longueur sur le rappel du nombre de syllabes contenues dans un GI est aussi appuyée par les données du test de Friedman appliqué à l'ensemble des GI, que ceux-ci contiennent un ou deux GR (voir Tableau II, p. 44). Ce test indique que plus le GI est long, moins bonne est la performance des sujets. En fonction du format des données utilisées pour l'application du test, un rang numérique élevé indique un grand nombre de rappels corrects. Notons que les résultats indiquent que la proportion de bons rappels est inversement proportionnelle à la longueur du GI en syllabes, ce qui tend à démontrer l'existence d'un effet de longueur affectant la mémoire immédiate pour les structures intonatives.

Cet effet de longueur s'est également révélé avec les données relatives à la distribution des bonnes réponses à travers l'ensemble des stimuli. Un calcul du pourcentage cumulatif de rappel associé à chaque longueur de GI permet d'établir à huit syllabes la limite de la mémoire immédiate pour les GI. C'est-à-dire que les GI de deux

à huit syllabes comptent pour un peu plus de 95 % du total des cas de reproduction correcte (96,27 %) et couvrent donc la portion significative des résultats. La contribution des groupes de neuf syllabes demeure marginale, avec moins de 5 % des cas, soit seulement 3,73 % des bonnes réponses. Ces proportions sont comparables aux résultats de l'Expérience I car ils présentent la même limite en terme de longueur de GI. Pour fin de comparaison avec la présente expérience, un nouveau calcul des données de l'Expérience I a été mené de façon à ne tenir compte que des GI de longueur semblable à ceux utilisés dans la tâche de mémoire (de deux à neuf syllabes). La Figure 26 (p. 46) illustre la correspondance entre les pourcentages cumulatifs des courbes de réponse au présent test et ceux de la production de GI en parole spontanée. Pour les GI de neuf syllabes ou moins, les deux courbes se rejoignent à environ huit syllabes, ce qui tend à confirmer que les GI que l'on retrouve en parole sont bien ceux qui sont le plus facilement mémorisés.

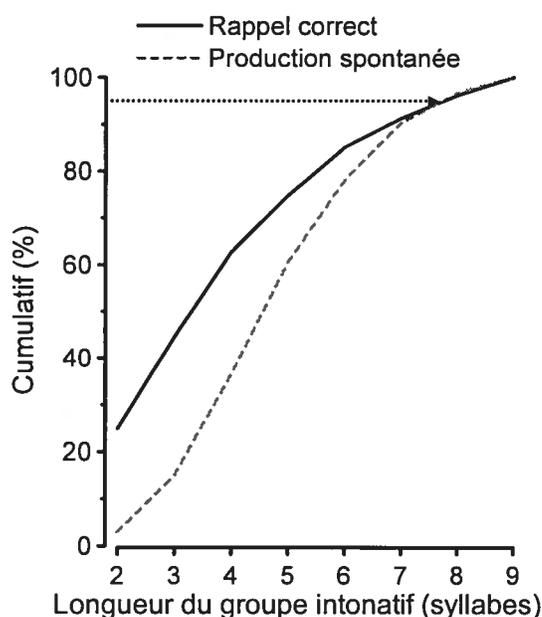


Figure 26 : % cumulatif en fonction du nombre de syllabes dans le GI en parole spontanée et en tâche de rappel. La limite dans les performances de rappel de GI correspond approximativement à la limite de longueur des GI en parole.

Cependant, la différence entre les deux courbes au niveau des GI plus courts indique que la fréquence relative des GI en parole se situe d'avantage vers des longueurs moyennes. Ces effets de fréquence d'occurrence ne sont pas associables aux conditions du test de rappel. Remarquons toutefois que les GI sont généralement composés d'un ou deux GR (Sityaev 1999) et que les GR contiennent en moyenne 3,5 syllabes (Dauer 1983, Martin 1999). Il est donc peu probable qu'une combinaison de GI et de GR donne lieu à des GI très courts (deux ou trois syllabes), alors qu'elle tend naturellement vers la création de GI contenant entre 3,5 et 7,1 syllabes. Les groupes très courts sont donc «mathématiquement» moins probables en parole spontanée malgré la facilité avec laquelle ils peuvent être rappelés. En somme, ces statistiques (ainsi que celles du Tableau III) indiquent qu'en considérant les GI de deux à neuf syllabes, le taux de rappel global attribuables aux GI de huit syllabes et moins (96,27 %) rejoint assez précisément la fréquence cumulative des GI de huit syllabes et moins en parole spontanée (94,89 %).

Enfin, notons que les interprétations ci-dessus s'appliquent à des sujets non musiciens et que la performance supérieure des sujets musiciens peut s'expliquer par leur entraînement à la mémorisation de rythmes et de mélodies. Des résultats analogues ont aussi été obtenus par Thompson, Schellenberg et Husain (2003), qui ont trouvé que les sujets ayant une importante formation musicale (plus de 15 ans) étaient plus habiles à extraire la prosodie d'un énoncé entendu pour la comparer à des séquences de tons synthétiques. Néanmoins, ces effets de formation en musique n'infirment aucunement la correspondance observée entre la limite générale de huit syllabes pour les GI en parole et la limite approximative de huit syllabes qui s'applique au rappel des schèmes intonatifs. Il est, par ailleurs, raisonnable de supposer que les structures prosodiques entendues qui sont difficilement reproduites n'auraient pas tendance à être apprises et utilisées en parole.

Toutefois, en reconnaissant que cet effet de limite peut provenir de contraintes mnémoniques, on admet qu'un principe indépendant de la syntaxe peut agir sur l'organisation prosodique. La prochaine expérience vise à démontrer que ce principe est un facteur de prédiction de la prosodie qui opère de façon indépendante de la syntaxe.

6. Expérience III

Les expériences décrites ci-dessus ont établi que la mémoire immédiate exerce une contrainte sur la longueur des GI qui peuvent être reproduits, ce qui aura nécessairement une influence sur les GI qui peuvent être utilisés en parole. L'objectif des deux expériences précédentes, rappelons-le, était de démontrer que la mémoire immédiate peut constituer un facteur sous-tendant la production des GI en parole spontanée, expliquant ainsi l'absence de correspondance entre les schèmes prosodiques et les constituants syntaxiques postulés en linguistique formelle. La présente expérience vise à illustrer comment les contraintes de la mémoire immédiate peuvent être utilisées afin de prédire la longueur des GI, et ce, indépendamment de la syntaxe. Pour ce faire, une tâche a été élaborée ayant pour but de déterminer laquelle de la contrainte de mémoire ou de la syntaxe prédira le mieux la production des GI dans des phrases contenant des constituants de longueur différente, mais dont la syntaxe demeure inchangée. Si la syntaxe prédit efficacement la production de groupes intonatifs, on pourra croire qu'elle représente le facteur dominant dans l'organisation du GI indépendamment des contraintes cognitives. D'autre part, si la contrainte mnémorique prédit mieux la production, on pourra penser que cette contrainte est un facteur dominant qui influence les structures prosodiques que l'on retrouve dans le langage.

6.1. Sujets

Les quarante-deux sujets de l'Expérience II ont aussi participé à l'Expérience III. On a conservé les données provenant de sujets ayant une formation en musique et/ou en danse car les risques de biais sont minimes vu le type de tâche demandé. On a cependant dû écarter les résultats de deux sujets, parce qu'ils n'avaient pas compris la tâche correctement, et ce, même après un rappel des directives. Il restait donc quarante sujets locuteurs natifs du français (21 hommes, 19 femmes) âgés entre 19 et 32 ans (moyenne : 24,63 ans).

6.2. Stimuli

Deux groupes de phrases ont été élaborés dans le but d'isoler l'effet de la longueur des différents constituants syntaxiques sur la production de GI. Elles ont été construites de façon à maintenir une structure syntaxique de base constante à l'intérieure de chaque groupe et à créer des groupes rythmiques de longueur inégale afin d'éviter les effets de symétrie. On a aussi évité d'utiliser des mots dont le nombre de syllabes risquait d'être modifié lors de la production par des phénomènes phonétiques tels que la liaison et la syncope. Dans ces phrases, le SN-sujet est toujours un nom propre, afin de permettre son accentuation finale.

Le premier groupe de stimuli consiste en quatre phrases de douze syllabes, dont la structure syntaxique est maintenue à SVO_{direct}Épithète (SN, SV, (SN, (SAdj))), et où la variable indépendante est l'emplacement relatif de la coupure syntaxique majeure entre le sujet (SN) et le verbe (SV) à l'intérieur de la phrase. La longueur de l'objet direct est gardée constante (sept syllabes) afin d'éviter les biais qui pourraient être dus aux variations de sa longueur. Ces phrases, présentées dans le Tableau IV (p. 50), ont été utilisées afin de trouver lequel de la présence d'une frontière syntaxique majeure ou des limites de la mémoire trouvées dans l'Expérience II prédirait le mieux la production du groupe intonatif

Le deuxième groupe de stimuli est composé de trois phrases de douze syllabes construites selon une structure syntaxique SVO_{direct} (SN, SV, (SN)), et où la variable indépendante est la longueur totale (en syllabes) du syntagme verbal (SV(SN)) exempt de frontière syntaxique majeure. Le sujet syntaxique (SN) de ces phrases était toujours un nom propre de trois à cinq syllabes. Des prétests ont démontré que les noms propres ont tendance à être produits dans un groupe intonatif à part, ce qui permet d'observer le GI du syntagme verbal. Ces stimuli ont été utilisés afin de voir si les longs syntagmes verbaux sont soumis à la contrainte mnémonique sur le nombre de syllabes établie lors de l'expérience précédente. Pour ce faire on a utilisé des longueurs de VP se rapprochant de la longueur maximale (huit syllabes) trouvée lors de l'Expérience I. Le Tableau V (p. 50) présente les phrases du deuxième groupe.

Tableau IV : *Phrases du premier groupe en fonction du nombre de syllabes avant et après la frontière syntaxique majeure.*

SN		SV
Luc (1 syll)	entretiendrait (4 syll)	les rosiers décoratifs.
5 syll		7 syll
Martin (2 syll)	apprécie (3 syll)	la musique instrumentale.
5 syll		7 syll
Dominic (3 syll)	ajoute (2 syll)	des brisures chocolatées.
5 syll		7 syll
Maximilien (4 syll)	vend (1 syll)	des chatons vermifugés.
5 syll		7 syll

Tableau V : *Phrases du deuxième groupe en fonction du nombre de syllabes dans les constituants.*

SN	SV
Alexandrina (5 syll)	regardait les candidats. (7 syll)
Élizabeth (4 syll)	psychanalysait les enfants. (8 syll)
Anabelle (3 syll)	investissait ses économies. (9 syll)

6.3. Procédure

Les stimuli ont été présentés visuellement à l'intérieur d'une cabine à l'épreuve du bruit et par le biais d'un écran LCD placé à 70 cm du sujet (± 10 cm). Les phrases y apparaissaient complètes, centrées, en caractères blancs (Arial Unicode gras de 45 points) sur fond noir. Aux phrases d'intérêt on avait ajouté d'autres phrases utilisées pour une autre expérience non-présentée ici. Chaque phrase était précédée d'un avertisseur visuel (*). La durée totale de la tâche ne dépassait pas 15 minutes. Le sujet devait lire à voix haute les instructions qui lui étaient présentées à l'écran avant que ne débute la présentation des phrases-stimuli. La tâche du sujet consistait, dans un premier temps, à lire la phrase qui lui était présentée à l'écran. Ensuite, la phrase disparaissait et le sujet devait répéter la phrase de mémoire. Seule la version « rappelée » faisait l'objet d'analyses. Deux exemples servant de pratique ont été utilisés au début de la présentation des stimuli, afin de s'assurer que le sujet avait bien compris la tâche à exécuter. Par ailleurs, si le sujet utilisait un ton récitatif durant les exemples, l'expérimentateur lui demandait de prononcer la phrase comme il le ferait spontanément en contexte amical. Ce correctif aux directives n'a été donné qu'une seule fois par sujet à environ 25 % des participants, les autres n'ayant pas besoin de ces directives supplémentaires. Deux sujets qui n'arrivaient pas à ajuster leur production n'ont pas reçu de directives supplémentaires mais ont plutôt été écartés de l'analyse.

L'expérimentateur contrôlait le déroulement de l'expérience à partir de l'extérieur de la cabine pour s'assurer que le rythme de présentation des stimuli était constant et adapté à la vitesse de réponse de chaque sujet. On a maintenu un contact audio entre l'intérieur et l'extérieur de la cabine en tout temps. Cette méthode permettait à l'expérimentateur de vérifier si les sujets lisaient et rappelait correctement les phrases données. Advenant une erreur impliquant un changement de longueur (en syllabes) d'un des constituants lors du rappel, le stimulus était resoumis au sujet une deuxième fois (rarement plus d'une fois), jusqu'à ce qu'il soit reproduit avec le bon nombre de syllabes. Les réponses des sujets ont été enregistrées numériquement à l'aide du même matériel que celui utilisé dans l'Expérience II et en suivant les mêmes ajustements.

6.4. Analyse

L'établissement des frontières de GI et le décompte des syllabes s'est fait selon la même méthode que celle utilisée lors de l'Expérience I.

6.5. Résultats

La fiabilité de l'analyse a été établie par une réanalyse de 15 % des résultats, ce qui a donné un taux de correspondance test-retest de 97,62 % (41/42).

Dans les phrases du premier groupe, la présence d'une coupure syntaxique majeure n'était pas un facteur stable de prédiction de la production du GI. La proportion des bonnes prédictions faites par la frontière syntaxique majeure varie entre 22,5 % et 82,5 %, et augmente en fonction du nombre de syllabes contenues dans le sujet (SN), comme on peut le voir au Tableau VI (p. 52). Un test non paramétrique (Q de Cochran) a démontré qu'il y avait une différence significative dans la proportion de bonne prédiction en fonction de la longueur du sujet ($Q = 33,81$, $n = 40$, *Sig. Asymp.* = 0,000). Autrement dit, plus le SN est long, meilleures sont les chances qu'il sera produit dans un seul groupe intonatif et pourra être prédit en fonction de la frontière syntaxique majeure.

Tableau VI : *Nombre de bonnes prédictions par la syntaxe (frontière syntaxique majeure = frontière de groupe intonatif) en %.*

Longueur du sujet (SN)	Prédiction correcte par la syntaxe (%)
1	22,5
2	62,5
3	77,5
4	82,5

En ce qui a trait au deuxième groupe de stimuli, il faut d'abord noter que deux réponses ont dû être retirées de l'analyse, car le sujet n'amorçait pas un nouveau GI avec le syntagme verbal comme supposé lors de la composition des stimuli. Les résultats obtenus indiquent une tendance inverse de ceux du premier groupe de stimuli, car les probabilités qu'un syntagme verbal soit produit en un seul GI diminuent en fonction de sa longueur, et ce, même si la structure syntaxique demeure inchangée.

En somme, les résultats confirment que les syntagmes verbaux dont la longueur en syllabes dépasse la longueur maximale des groupes intonatifs produits en parole spontanée (huit syllabes) sont rarement produits en un seul GI. À l'opposée, les syntagmes plus courts sont facilement et fréquemment contenus à l'intérieur d'un seul GI. Le Tableau VII (p. 53) présente les résultats obtenus pour le deuxième groupe de stimuli. La contrainte de longueur est notée par une baisse marquée de la production de GI de neuf syllabes par rapport aux autres longueurs de GI. Ces données sont aussi appuyées par le résultat d'un test du Q de Cochran qui démontre qu'il y a une différence significative entre les proportions de bonne prédiction selon la longueur du SV ($Q = 28,692$, $n = 40$, $Sig. Asymp. = 0,000$).

Tableau VII : *Nombre de bonnes prédictions par la syntaxe (absence de frontière syntaxique = absence de frontière de groupe intonatif) en %.*

Longueur de l'objet (SN(SAdj))	Production d'un seul GI (%)
7	67,5
8	58,97
9	12,82

6.6. Discussion

Les données présentées dans les Tableaux VI et VII montrent que la production des GI varie en fonction du nombre de syllabes qu'ils contiennent, et ce, malgré une structure syntaxique constante. Plus la longueur du syntagme présenté se rapproche de la longueur moyenne des GI trouvée lors de l'Expérience I, plus il est probable d'obtenir un seul GI, peu importe la structure syntaxique. Un effet de limite est de plus confirmé par la baisse importante de la production des GI de neuf syllabes par rapport à ceux de huit syllabes. Cette limite est comparable à celle trouvée dans l'Expérience I sur la longueur maximale des GI en parole spontanée et à celle de l'Expérience II sur les limites de la mémoire immédiate des GI. Ces résultats indiquent donc que la longueur (en syllabes) d'un syntagme agit comme facteur déterminant de la production du GI, et la structure syntaxique apparaît comme un facteur secondaire. En fait, la syntaxe ne servirait pas à prédire la longueur des GI produits, mais permettrait de définir où pourrait avoir lieu une coupure, advenant que le nombre de syllabes oblige à une segmentation de la chaîne parlée.

7. Discussion générale et conclusions

Pour tester l'hypothèse générale décrite à la section 3 selon laquelle il existe un lien entre les GI produits en parole spontanée et la mémoire immédiate il fallait 1) connaître la longueur maximale des GI en parole spontanée, 2) voir si la même longueur maximale s'appliquait aussi au rappel du nombre de syllabes dans des GI entendus et 3) tester l'applicabilité de ladite contrainte de longueur pour prédire la production de GI en parole dirigée. Les Expériences I, II et III décrites précédemment ont servi à vérifier cette hypothèse par l'entremise de trois sous hypothèses.

L'Expérience I visait à établir des statistiques sur la longueur des GI en parole spontanée. En fonction des articles publiés sur le nombre de GR par GI (Sityaev 1999) et la longueur moyenne des GR en parole (Dauer 1983, Martin 1999), on prévoyait obtenir une majorité de GI contenant entre 3,5 et 7,1 syllabes. Les résultats obtenus confirment en partie cette prédiction dans la mesure où 85,34 % des GI observés en parole comptent entre trois et sept syllabes. On a aussi établie à huit syllabes la longueur maximale des GI (comptant pour 95 % des cas).

L'Expérience II avait pour but de voir si cette limite de huit syllabes par GI trouvée en parole s'appliquait aussi au rappel des longueurs de GI. En se basant sur les résultats obtenus par d'autres auteurs sur le lien entre la mémoire et les GR, on a émis l'hypothèse qu'une contrainte de longueur s'appliquerait aussi bien à la production des GI en parole qu'au rappel exact de ces structures. Pour vérifier cette hypothèse il fallait d'abord établir l'effet indépendant qu'avait la présence d'une frontière de GI sur le rappel. Les données expérimentales ont confirmé que la frontière de GI avait une influence sur le rappel du nombre de syllabes. On a par la suite testé l'influence de la longueur du GI sur le rappel du nombre de syllabes le composant. Les résultats indiquent que les GI de 8 et 9 syllabes étaient significativement plus difficiles à rappeler que les GI de 5 et 6 syllabes, ce qui tend à démontrer la présence d'une contrainte du même ordre que celle trouvée avec l'Expérience I. On a de plus fait une analyse détaillée de l'ensemble des GI présentés qui a permis de confirmer que la contrainte agissant sur la mémoire immédiate limitait effectivement à une longueur approximative de huit syllabes les GI pouvant être rappelés efficacement. Ces résultats appuient de

façon générale l'hypothèse spécifique à l'Expérience II voulant que la même limite de longueur s'applique à la fois au rappel du nombre de syllabes entendues et à la production de GI en parole spontanée.

En somme, en produisant de la parole spontanée, un locuteur tend à limiter ses GI (et ses GR) à une certaine longueur et aucune théorie actuelle n'offre une explication de cette contrainte. Dans ce travail on a avancé l'hypothèse que la mémoire immédiate pourrait contraindre les schèmes prosodiques dans la parole. Ce lien semble logique car tout schème difficile à reproduire n'aurait pas tendance à être appris puis utilisé en parole. De plus, comme la reproduction de séquences entendues implique la mémoire, l'existence d'un lien entre les limites du rappel et les limites sur les schèmes prosodiques en parole semble motivé et valable. Toutefois, ce type de lien confronte l'idée générale voulant que la syntaxe soit déterminante dans l'organisation de la prosodie. L'Expérience III visait à démontrer qu'au contraire l'effet de longueur tirant son origine de processus mnémoniques pouvait être un facteur important sous-tendant la production de GI lorsque ceux-ci atteignent une certaine longueur limite.

L'objectif de l'Expérience III était de vérifier l'applicabilité de la contrainte de longueur à la prédiction des GI d'énoncés français-québécois. On a émis l'hypothèse que la contrainte de longueur des GI mise au jour dans les Expériences I et II pourrait mieux prédire la production des GI que la syntaxe dans une tâche de production dirigée. On a soumis aux sujets deux batteries d'énoncés dans lesquelles on a fait varier soit l'emplacement de la frontière syntaxique majeure, soit la longueur du VP au-delà de huit syllabes. Les résultats obtenus ont permis de confirmer l'hypothèse de deux façons. D'une part, on a trouvé que la fiabilité des prédictions basées sur la syntaxe est dépendante de la longueur du NP sujet. Selon ces résultats, plus le NP sujet est long et plus les prédictions à base de syntaxe sont fiables. D'autre part, les sujets n'ont pas produit (dans des proportions significatives) de GI de plus de huit syllabes, et ce, même si la structure syntaxique de l'énoncé permettait théoriquement la production de constituants prosodiques plus long. Ceci démontre que la longueur des constituants (en syllabes) représente un facteur important, sinon déterminant de prédiction de la prosodie produite.

Les trois expériences précédentes ont donc permis de confirmer l'hypothèse générale selon laquelle la mémoire immédiate restreint la production des structures prosodiques de la parole. De plus, en tenant compte aussi des résultats de Boucher (à paraître) sur l'existence d'un lien entre la mémoire et les GR de la parole spontanée, on peut avancer que la prosodie serait le reflet de contraintes d'ordre cognitif indépendantes de la « compétence linguistique », tel que le suppose l'idée reçue à l'heure actuelle et qui a été fortement intégrée dans les perspectives de Chomsky (1967).

Bien que l'existence d'une contrainte régissant à la fois la mémoire immédiate des GI et leur production en parole ait été démontrée, il reste encore à déterminer le rôle précis que joue l'intonation dans la compréhension du langage. En effet, l'intonation (GI) est systématiquement présente dans chaque énoncé produit oralement, et ce, même si ces schèmes imposent une charge mnémonique comme on l'a démontré à l'Expérience II. Quant à savoir pourquoi un phénomène imposant une telle charge mnémonique ne serait pas disparu au cours de l'évolution des langues, on pourrait penser que ces schèmes sont indispensables à l'interprétation du message oral.

8. Bibliographie

- Aldridge, J. W. (1978). Levels of processing in speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, 4(1), 164-177.
- Aldridge, J. W. (1981). Levels of processing in speech production. *Journal of Experimental Psychology: Human Perception and Performance*, 7(2), 388-407.
- Baddeley, A. (2000). Short-term and working memory. In E. Tulving et F. I. M. Craik (Eds.), *The Oxford Handbook of Memory*. (pp.77-92). Oxford: Oxford University Press.
- Beckman, M. E. et Pierrehumbert, J. (1986). Intonational structure in Japanese and English, *Phonology Yearbook*, 3, 255-309.
- Boucher, V. J. (1994). Alphabet-related biases in psycholinguistic enquiries: considerations for direct theories of speech production and perception. *Journal of Phonetics*, 22, 1-18.
- Boucher, V. J. (à paraître). On the function of stress rhythms in speech: Evidence of a link with grouping effects on serial memory. *Language and Speech*.
- Boucher, V. J. (2006). Intrinsic factors of cyclical motion in speech articulators: Reappraising postulates of segmental commands in motor-control theories. *Manuscrit*.
- Boucher, V. J., et Gilbert, A. C. (2005). Constraints on immediate memory for prosody: Evidence of a correspondence with the prosodic organization of speech. *Proceedings of the XXVII Annual Conference of the Cognitive Science Society*, 328-330.
- Byrd, D. et Saltzman, E. (2003). The elastic phase: Modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics*, 31, 149-180.
- Byrne, B., et Ledez, J. (1986). Phonological awareness in reading-disabled adults. *Australian Journal of Psychology*, 35, 185-197.
- Chomsky, N. (1967). The formal nature of language. In E. H. Lenneberg (Ed.), *Biological Foundations of Language* (pp. 397-442). New York: John Wiley & Sons.
- Cole, R. A. (1973). Listening for mispronunciations: A measure of what we hear during speech. *Perception and Psychophysics*, 14 153-156.
- Cooper, W. E. et Sorensen, J. M. (1981). *Fundamental frequency in sentence production*. New York, Springer Verlag.
- Cruttenden, A. (1986). *INTONATION*. Cambridge: Cambridge University Press.

- Cutler, A. (1994). The perception of rhythm in language. *Cognition*, 50, 79-81.
- Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics*, 11, 51-62.
- Delattre, P. (1966a). A comparison of syllable length conditioning among languages. *International Review of Applied Linguistics*, 4, 183-198.
- Delattre, P. (1966b). Les dix intonations de base du français. *The French Review*, 40(1), 1-14.
- DeZitter, A. (1992). The perceptual salience of melodic variation: contour shape and peak height. *Journal of Phonetics*, 20, 181-188.
- DiCristo, A., et Hirst, D. J. (1986). Modelling French micromelody: Analysis and synthesis. *Phonetica*, 43, 11-30.
- Dommergues, J.-Y., et Grosjean, F. (1981). Performance structures in the recall of sentences. *Memory and Cognition*, 9(5), 478-486.
- Ehri, L., et Wilce, L. (1980). The influence of orthography on readers' conceptualization of the phonemic structure of words. *Applied Psycholinguistics*, 1, 371-385.
- Fant, G., Kruckensberg, A., et Nord, L. (1991). Durational correlates of stress in Swedish, French and English. *Journal of Phonetics*, 19, 351-365.
- Fletcher, J. (1991). Rhythm and final lengthening in French. *Journal of Phonetics*, 19, 193-212.
- Fougeron, C. (1999). Articulation of Prosodic Phrasing in French. *Proceedings of the 14th International Congress of Phonetic Science*, 675-678.
- Frankish, C. (1989). Perceptual organization and precategorical acoustic storage. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 15(3), 469-479.
- Frankish, C. (1995). Intonation and auditory grouping in immediate serial recall. *Applied Cognitive Psychology*, 9, 5-22.
- Glanzer, M. (1976). Intonation grouping and related words in free recall. *Journal of Verbal Learning and Verbal Behavior*, 15, 85-92.
- Gussenhoven, C., et Rietveld, A. C. M. (1992). Intonation contours, prosodic structure and preboundary lengthening. *Journal of Phonetics*, 20, 283-303.
- Hirst, D. (1993). Detaching intonational phrases from syntactic structure. *Linguistic Inquiry*, 24(4), 781-788.

- Hitch, G. J., Burgess, N., Towse, J. N., et Culpin, V. (1996). Temporal grouping effects in immediate recall: A working memory analysis. *The Quarterly Journal of Experimental Psychology*, 49A(1), 116-139.
- Kenning, M.-M. (1983). The tones of English and French. *Journal of Phonetics*, 13(1), 32-48.
- Lacheret-Dujour, A., Martel, K., et Vivier, J. (1999). Iconic aspects of prosody in French first language acquisition. *Proceedings of the 14th International Congress of Phonetic Science*, 2185-2188.
- Ladd, D. R. (1983). Peak features and overall slope. In A. Cutler et D. R. Ladd (Eds.), *Prosody: Models and Measurements* (pp. 39-52). New-York: Springer-Verlag.
- Ladefoged, P. (1984). The limits of biological explanations in phonetics. *UCLA Working Papers in Phonetic*, 59, 1-10.
- Leonard, L. B. (1974). The role of intonation in the recall of various linguistic stimuli. *Language and Speech*, 16, 327-335.
- Mann, V. (1986). Phonological awareness: the role of reading experience. *Haskins Laboratories: Status Report on Speech Research*, SR-85, 1-22.
- Marcus, S. M. (1981). Acoustic determinants of perceptual center (P-center) location. *Perception & Psychophysics*, 30, 247-256.
- Martin, P. (1987). Prosodic and rhythmic structures in French. *Linguistics*, 25, 925-949.
- Martin, P. (1991). A propos du statut théorique de la structure prosodique de la phrase. *Verbum*, 14(2-4), 307-310.
- Martin, P. (1999). Intonation of spontaneous speech in French. *Proceedings of the 14th International Congress of Phonetic Sciences*, 17-20.
- Maybery, M. T., Permentier, F. B. R., et Jones, D. M. (2002). Grouping of list items reflected in the timing of recall: implications for models of serial verbal memory. *Journal of Memory and Language*, 47(3), 360-385.
- Millotte, S. (2005). *Le rôle de la prosodie dans le traitement syntaxique adulte et l'acquisition de la syntaxe*, Thèse de doctorat.
- Miyake, A., et Shah, P. (1999). MODELS OF WORKING MEMORY. Cambridge : Cambridge University Press.
- Monnin, P., et Grosjean, F. (1993). Les structures de performance en Français: Caractéristiques et prédiction. *L'Année Psychologique*, 93, 9-30.
- Morley, M. E. (1957). *THE DEVELOPMENT AND DISORDERS OF SPEECH IN CHILHOOD*. London : Livingston.

- Morais, J., Cary, L., Alégria, J., et Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition*, 7, 323-331.
- Nooteboom, S. (1997). The prosody of speech: Melody and rhythm. In W. J. Hardcastle & J. Laver (Eds.), *Handbook of Phonetic Sciences* (pp. 640-673). Oxford: Blackwell Publishers.
- Read, C., Zhang, Y., Nie, H., et Ding, B. (1986). The ability to manipulate speech sounds depends on knowing alphabetic writing. *Cognition*, 24, 31-44.
- Reeves, C., Schauder, A. R., et Morris, R. K. (2000). Stress grouping improves performance on an immediate serial list recall task. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 26(6), 1638-1654.
- Rosner, B. S., Grabe, E., Nicholson, H., Owen, K., et Keane, E. (2003). Prosody, memory load, and memory for speech. *Oxford University Working Papers in Linguistics, Phonology and Phonetics*, 8, 85-102.
- Saito, S. (2001). The phonological loop and memory for rhythms: An individual differences approach. *Memory*, 9(4-5-6), 313-322.
- Scholes, R. H. (1971). *ACOUSTIC CUES FOR CONSTITUENT STRUCTURE*. The Hague, The Netherlands.: Mouton & Co.
- Selkirk, E. O. (1984). *PHONOLOGY AND SYNTAX: THE RELATIONSHIP BETWEEN SOUND AND STRUCTURE* Cambridge: MIT Press.
- Sityaev, D. (1999). Distribution of pitch accents in intonation groups across intonational styles. *Proceedings of the 14th International Congress of Phonetic Sciences*, 1505-1508.
- Springer, S. P. (1973). Memory for linguistic and non-linguistic dimensions of the same acoustic stimulus. *Journal of Experimental Psychology*, 101, 159-163.
- Thomas, M. G., et Hutchens, T. A. (1990). The effects of vocal intonation in digit span testing. *Journal of Psychoeducational Assessment*, 8, 150-154.
- Thompson, W. F., Schellenberg, E. G. et Husain G. (2003). Perceiving Prosody in Speech: Effects of Music Lessons. *Annals of the New York Academy of Sciences* 530-532.
- Vaissière, J. (sous presse). Sur les universaux de substance prosodique. In S. Wauquier (Ed.), *Les Universaux Sonores*. Nantes: Presses Universitaires de Rennes.
- Vaissière, J. (1995). Phonetic explanation for cross-linguistic prosodic similarities. *Phonetica*, 52, 123-130.
- Wolf, C. G. (1977). The processing of fundamental frequency in a dichotic matching task. *Brain and Language*, 4, 70-77.

Annexe

Tableau VIII : Moyenne et écart par rapport à la moyenne (z score) pour les différentes longueurs de GI produites par les sujets à l'expérience I. Notons que les GI de plus de neuf syllabes ont été regroupés car ils n'ont pas été produits par tous les sujets, de même que les groupes de moins de trois syllabes.

Sujet	Longueur du GI en syllabes																	
	< 3	3	4	5	6	7	8	9	9 <									
S1	1,74	-0,62	6,96	-0,86	22,61	0,11	24,35	0,3	17,39	-0,03	14,78	0,63	7,83	0,54	0,87	-0,82	3,48	0,75
S2	2,33	-0,26	13,95	0,73	14,73	-1,61	33,33	1,9	13,18	-0,87	10,08	-0,55	6,98	0,1	3,1	-0,08	2,33	0,24
S3	2,45	-0,19	17,18	1,47	22,09	0	25,15	0,45	19,63	0,41	7,98	-1,08	4,29	-1,31	1,23	-0,7	0	-0,78
S4	4,35	0,97	16,3	1,27	20,65	-0,32	21,74	-0,16	10,87	-1,33	15,22	0,74	8,7	1	2,17	-0,39	0	-0,78
S5	3,51	0,46	6,14	-1,05	21,93	-0,04	27,19	0,81	13,16	-0,88	14,04	0,44	8,77	1,04	3,51	0,05	1,75	-0,01
S6	0	-1,68	5,88	-1,11	32,35	2,23	17,65	-0,89	29,41	2,37	8,82	-0,87	5,88	-0,48	0	-1,11	0	-0,78
S7	5,56	1,71	17,36	1,51	20,83	-0,28	19,44	-0,57	16,67	-0,18	7,64	-1,17	6,94	0,08	5,56	0,73	0	-0,78
S8	4,05	0,79	13,51	0,63	18,24	-0,84	22,3	-0,06	19,59	0,41	15,54	0,82	4,05	-1,43	2,03	-0,44	0,68	-0,48
S9	1,52	-0,76	4,55	-1,41	25,76	0,8	27,27	0,82	16,67	-0,18	16,67	1,11	6,06	-0,38	1,52	-0,61	0	-0,78
S10	0	-1,68	11,11	0,08	22,22	0,03	11,11	-2,06	14,81	-0,55	11,11	-0,29	11,11	2,26	11,11	2,579	7,41	2,47
S11	4,62	1,13	7,69	-0,7	18,46	-0,8	20	-0,47	20	0,49	16,92	1,17	6,15	-0,33	4,62	0,42	1,54	-0,1
S12	2,29	-0,29	6,87	-0,88	22,9	0,17	18,32	-0,77	17,56	0	17,56	1,33	5,34	-0,76	8,4	1,68	0,76	-0,44
S13	4,29	0,93	8,57	-0,49	15,71	-1,39	25,71	0,54	25,71	1,63	7,14	-1,29	5,71	-0,56	1,43	-0,64	5,71	1,73
S14	1,82	-0,57	13,64	0,66	24,55	0,53	29,09	1,15	17,27	-0,06	6,36	-1,49	5,45	-0,7	1,82	-0,51	0	-0,78
S15	2,86	0,06	11,43	0,16	28,57	1,41	17,14	-0,98	11,43	-1,22	14,29	0,51	8,57	0,93	2,86	-0,16	2,86	0,48
total	2,76	1,64	10,74	4,39	22,11	4,58	22,65	5,62	17,56	5,01	12,28	3,97	6,79	1,91	3,35	3,01	1,77	2,28



Handwritten text at the bottom right of the page, possibly a signature or date, which is mostly illegible due to fading.