

# Indices de frontières lexicales avec des pseudo-mots : une étude préliminaire

Alain Content<sup>1,2</sup>, Odile Bagou<sup>2</sup>, Uli Frauenfelder<sup>2</sup> & Cécile Fougeron<sup>3</sup>

1 Laboratoire de Psychologie Expérimentale – Université libre de Bruxelles – Av. F. Roosevelt 50, B-1050 Bruxelles

2 Laboratoire de Psycholinguistique Expérimentale – Université de Genève – 40 Bd du Pont d'Arve, CH-1205 Genève

3 Laboratoire de Phonétique et Phonologie – CNRS/Université de Paris 3 – 19 rue des Bernardins, F-75005 Paris  
alain.content@ulb.ac.be; odile.bagou@pse.unige.ch; ulrich.frauenfelder@pse.unige.ch; cecile.fougeron@univ-paris3.fr

## ABSTRACT

This study examines the durational and spectral correlates of word boundaries with pseudo-word sequences. The productions from one female native French speaker were extracted from short sentence contexts that induced two lexical segmentations of the same phonemic strings, either a word-final enchaînement consonant (CCV<sub>1</sub>C<sub>p</sub>#V<sub>2</sub>CV) or a word-initial consonant (CCV<sub>1</sub>#C<sub>p</sub>V<sub>2</sub>CV). Longer durations and lower F<sub>1</sub> were observed for V<sub>1</sub> in CCV<sub>1</sub>C<sub>p</sub>#V<sub>2</sub>CV. Similarly, longer durations were found for C<sub>p</sub> and V<sub>2</sub> when corresponding to word onset. Thus, phonetic cues to lexical segmentation are available in the signal, even for non-word constituents. Although the perceptual influence of these subtle cues remains to be demonstrated, the findings are consistent with a syllable-based segmentation heuristic that could be guided by such lexical structure cues to avoid missegmentations.

## 1. INTRODUCTION

Un problème majeur de la reconnaissance des mots dans la parole continue est de déterminer la segmentation lexicale. Nous avons proposé précédemment [1, 2] que les débuts de syllabes constituent des points d'alignement privilégiés pour la recherche lexicale et donc, que le locuteur (francophone) présuppose que chaque début de syllabe représente un début de mot. Ce principe ne peut toutefois pas fonctionner comme une règle absolue, mais comme une heuristique probabiliste, puisque l'alignement entre syllabe et mot n'est pas toujours vérifié. La validité de l'hypothèse dépend donc étroitement de la fréquence des cas de désalignement.

Plusieurs phénomènes phonologiques courants en français pourraient rendre cette hypothèse de segmentation aux attaques syllabiques moins efficace, voire peu plausible. Les descriptions usuelles de l'enchaînement (*bel ami*) et de la liaison (*des amis*) supposent en effet une resyllabation, de sorte que l'alignement entre débuts de syllabes et débuts de mots serait compromis. C'est la raison qui nous a conduits à examiner de manière approfondie les caractéristiques acoustiques et articulatoires des segments aux frontières de mots.

Plusieurs études phonétiques indiquent l'existence d'indices acoustiques différents selon la nature de la frontière. Ainsi, une étude de Krakow [3] portant sur l'anglais met en évidence des différences phonétiques entre consonnes initiales et finales (*home E* vs. *hoe me*) mais ne permet pas de distinguer consonnes initiales de mots et initiales syllabiques (*hoe me* vs. *homey*). De la même manière, une étude de Quené [4] montre que les propriétés phonétiques des consonnes initiales de mots

sont utilisées par les auditeurs hollandais natifs pour segmenter des séquences ambiguës.

En français, Dumay et al [5] ont comparé des suites homophoniques variant par la frontière lexicale (ex. /tæteu/, extrait de *tante roublarde* ou de *temps troublant*). Ils mettent en évidence un allongement de la première voyelle en condition non-finale, et de la consonne à l'initiale de mot. Cependant, ces effets n'apparaissent que pour les groupes consonantiques OBLI, associant une obstruente à une liquide, et non pour les groupes SOB, constitués d'un /s/ et d'une occlusive.

Fougeron et al. [6, 7] ont poursuivi cette recherche en examinant des consonnes intervocaliques simples. Les analyses portent sur la consonne et les deux voyelles adjacentes (pré- et post-consonantiques), dans trois conditions: frontière lexicale (*cas légal*), enchaînement (*cale égale*) et frontière syllabique (*qualégal*). Les analyses mettent en évidence des différences significatives dans la durée de la consonne, plus longue en position initiale de mot ou de syllabe que dans l'enchaînement. Par ailleurs, la durée de la première voyelle est augmentée dans les conditions correspondant à deux mots par rapport à la condition de frontière syllabique. Ces résultats suggèrent donc que la structure lexicale de l'énoncé laisse des traces dans la forme de surface qui est réalisée, fournissant par conséquent des indices potentiels pour distinguer, par exemple, la réalisation d'une frontière syllabique et celle d'un enchaînement.

L'objet de la présente étude est d'approfondir l'analyse des indices acoustiques aux frontières de mots dans le cas particulier de l'enchaînement. Dans cette perspective, les caractéristiques acoustiques des consonnes d'enchaînement sont systématiquement comparées à celles de consonnes initiales de mot. Sur le plan méthodologique, l'originalité de ce travail provient de l'utilisation de séquences phonémiques sans signification et composées de pseudo-mots. La segmentation "pseudo-lexicale" correspondant à l'enchaînement ou à l'initiale de mot (ex.  $\epsilon\epsilon u\#ifa$  vs.  $\epsilon\epsilon u\#kifa$ ) est induite par le contexte d'apparition des pseudomots.

L'usage de pseudo-mots présente plusieurs intérêts. D'une part, il rend possible un examen plus complet des effets éventuels en variant systématiquement la consonne pivot mais également son environnement, procédure que les limitations du lexique français excluent avec de vrais mots. Il permet d'examiner les effets de la segmentation lexicale sur un domaine plus large, soit la totalité de la séquence, puisque les deux suites de phonèmes comparées sont strictement identiques. De plus, l'emploi de pseudo-mots évite l'écueil des variations d'acceptabilité

syntactique et sémantique. Enfin, la nouveauté des pseudo-mots aux yeux des locuteurs, permet d'argumenter en faveur d'un marquage "pur" des frontières, éliminant une interprétation alternative en termes de familiarité des formes lexicales engagées.

La présente étude est, à notre connaissance, une première tentative en ce sens. Elle vise essentiellement à établir si des caractéristiques phonétiques systématiques différencient les consonnes finales enchaînées des consonnes initiales. Elle permet aussi de savoir si les segments environnants offrent également des indices phonétiques de la structure lexicale qui seraient susceptibles d'être exploités par les auditeurs dans les processus de segmentation.

## 2. MÉTHODE

Trente-deux pseudo-mots trisyllabiques de structure CCV.CV.CV ont été construits puis insérés dans des paragraphes de façon à induire la réalisation de suites correspondant à la structure CCV#CVCV ou à la structure CCVC#VCV. La consonne à l'attaque de la 2<sup>e</sup> syllabe, qui apparaît tantôt à la finale du premier non-mot, tantôt à l'attaque du second, sera désignée dans la suite comme *consonne pivot* (C<sub>p</sub>). Nous désignerons les voyelles par leur indice syllabique, V<sub>1</sub>, V<sub>2</sub>, et V<sub>3</sub>, par CC le groupe initial, et par C<sub>3</sub> la consonne de la dernière syllabe.

### 2.1. Construction des pseudo-mots

Quatre séries de huit pseudo-mots ont été établies en variant la sonorité (+*son* ou -*son*) de la consonne pivot et celle de la consonne C<sub>3</sub> (Cf. Table 1). Les consonnes considérées comme -*son* sont les obstruantes /f s p k/ alors que les consonnes /l ε m n/ sont +*son*. Dans chaque série, les groupes consonantiques initiaux /be bl εε el ke kl pe pl/ apparaissent chacun une fois. Enfin, la voyelle V<sub>2</sub> est également contrebalancée: les quatre voyelles /a o i y/ apparaissent exactement deux fois par série. Les deux autres voyelles, V<sub>1</sub> et V<sub>3</sub>, ont été assignées de manière semi-aléatoire, afin d'obtenir une distribution équilibrée sur l'ensemble du matériel.

**Table 1** : Présentation des pseudo-mots utilisés.

-son -son	-son +son	+son -son	+son +son
bεoko su	beakyeo	beølyfεe	beilamø
blyfipee	blofynø	blenapy	blieomi
εεukifa	εεεekaly	εεelikae	εεønyeu
elupofεe	eløpilu	elemoki	elumyεo
keøpysi	keεεpaεo	keumifε	kealonu
kleεfakae	klefomae	kleεεapi	kleεino
peøsyπεe	peεεsine	peinosy	peomale
plaesaku	plisoma	pleεεysa	planile

### 2.2. Paragraphes d'induction

En vue d'obtenir les deux segmentations lexicales d'intérêt (εεu#kifa vs. εεuk#ifa), seize paragraphes ont été construits pour induire la production de la frontière. Comme l'indiquent les deux exemples ci-dessous, chacun comportait trois parties: une phrase inductrice, une question et une reprise assertive.

La phrase inductrice fournissait un contexte interprétatif et introduisait les deux pseudo-mots constitutifs de la séquence trisyllabique, comme s'il s'agissait respectivement d'un substantif (*un grou|un grouque*) et d'un qualificatif précédé d'un ou plusieurs adverbes (*assez bien kifa|ifa*). La phrase interrogative introduisait la première mention de la séquence en position post-verbale non finale. La reprise assertive fournissait une seconde mention de la séquence, également en position post-verbale non finale. Le focus de la réponse négative portait soit sur l'action (*Vous avez acheté...? Non, on a gagné...*), soit sur l'agent (*Non, c'est elle qui...*), de telle façon que l'objet soit toujours en position post-focale. Les structures utilisées garantissaient l'absence d'une frontière prosodique entre les deux membres des séquences étudiées.

IND *Au parc, j'ai trouvé un grou carrément à l'ombre et un grou assez bien kifa.*

Q-M<sub>1</sub> *Tu as dormi sur le grou kifa pendant la matinée?*

A-M<sub>2</sub> *Non, je suis allé m'asseoir sur le grou kifa pendant la matinée.*

IND *Pour les vacances, on hésitait entre un grouque très touristique et un grouque plus ifa.*

Q-M<sub>1</sub> *Vous avez acheté le grouque ifa pour trois semaines?*

A-M<sub>2</sub> *Non, on a gagné le grouque ifa pour trois semaines.*

Chacun des pseudo-mots a été inséré dans quatre paragraphes différents afin d'obtenir deux exemplaires distincts pour chaque segmentation (CCV#CVCV et CCVC#VCV). Deux blocs de 64 paragraphes ont été ensuite constitués, chacun comprenant deux exemplaires de chaque pseudo-mot dans l'une des deux segmentations.

Un bloc comptait une moitié de pseudo-mots avec la segmentation CCV#CVCV et l'autre moitié avec la segmentation CCVC#VCV, et cet arrangement était inversé dans l'autre bloc. Cette répartition présentait deux avantages. D'une part, les différentes versions d'un même pseudo-mot étaient distribuées sur l'ensemble de la liste, ce qui permettait d'éviter des effets de contamination liés aux variations d'attention ou de débit au cours de la session. D'autre part, apparaissant dans des blocs distincts, les segmentations différentes d'une même séquence étaient généralement éloignées l'une de l'autre. Masquée par les répétitions multiples de la même séquence dans le paragraphe de contexte proposé, la manipulation de frontière était ainsi moins apparente.

Suite à une lecture préalable pour prendre connaissance du matériel et corriger d'éventuelles erreurs de prononciation des pseudo-mots, une locutrice francophone belge ignorant les objectifs de l'étude a enregistré, en une seule séance, les deux blocs de paragraphes.

Les deux mentions de chaque pseudo-mot ainsi que la syllabe finale du mot précédent et la syllabe initiale du mot suivant, ont été extraites de l'enregistrement. Ces séquences ont été ultérieurement segmentées et analysées sous PRAAT par l'un des auteurs ne disposant d'aucune information relative au contexte de production d'origine. La durée acoustique des segments CC, V<sub>1</sub>, C<sub>p</sub>, V<sub>2</sub>, C<sub>3</sub> et

$V_3$ , la fréquence fondamentale ainsi que les caractéristiques spectrales des voyelles à leur point central ont été mesurées et comparées dans chacune des conditions de frontière établies.

### 3. RÉSULTATS

#### 3.1. Durées

Des analyses de variance à mesures répétées ont été réalisées sur la durée totale des séquences et sur la durée de chaque segment, en fonction de la mention (1<sup>ère</sup> : question; 2<sup>ème</sup> : assertion), de l'exemplaire (1<sup>ère</sup> ou 2<sup>ème</sup> répétition,  $R_1$  vs  $R_2$ ), et de la segmentation lexicale induite. Les séquences s'avèrent plus longues dans les phrases assertives que dans les phrases interrogatives (respectivement, 564.5 vs. 547.7 ms.,  $F[1, 31] = 24.8^1$ ), et le second exemplaire est en moyenne plus court que le premier (561 vs 551 ms.,  $F[1, 31] = 8.4$ ). L'effet du contexte pourrait être dû à des variations locales de débit, en fin de paragraphe. Il apparaît pour plusieurs segments (CC:  $F[1, 31] = 14.7$ ;  $V_1$ :  $F[1, 31] = 5.5$ ;  $C_p$ :  $F[1, 31] = 15.1$ ;  $C_3$ :  $F[1, 31] = 52.4$ ).

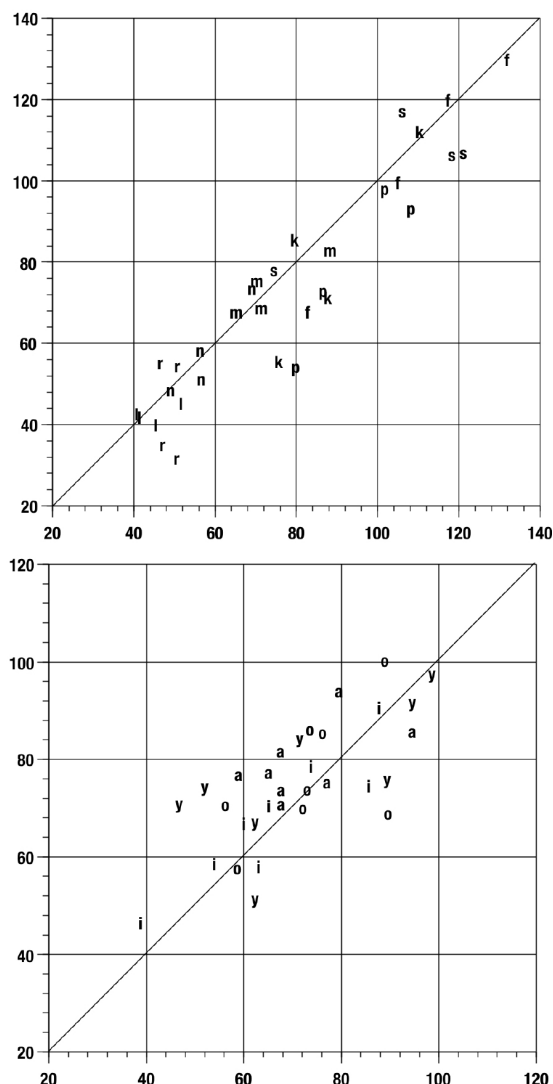
**Table 2 :** Durées des segments et de la séquence, selon l'exemplaire, la segmentation et le contexte de phrase.

	CC	$V_1$	$C_p$	$V_2$	$C_3$	$V_3$	Tot
CCV- $R_1$	133.4	79.9	78.4	70.9	92.4	94.4	549.5
CCV- $R_2$	133.0	77.9	77.1	71.2	89.4	95.3	544.0
CCVC- $R_1$	131.1	84.6	73.6	75.0	92.1	99.0	555.3
CCVC- $R_2$	125.3	83.2	73.1	75.8	89.4	95.5	542.1
Mention 1	130.7	81.4	75.5	73.2	90.8	96.0	547.7
CCV- $R_1$	138.1	77.2	80.2	71.7	100.2	99.0	566.4
CCV- $R_2$	136.0	74.2	81.9	72.5	97.9	97.4	559.9
CCVC- $R_1$	138.4	79.5	82.6	70.0	101.9	101.0	573.4
CCVC- $R_2$	137.3	79.0	79.0	69.9	100.4	93.1	558.6
Mention 2	137.4	77.5	80.9	71.0	100.1	97.6	564.5

Un effet significatif de la segmentation lexicale induite est observé pour la première voyelle ( $F[1, 31] = 5.7$ ), qui est allongée en position non-finale (77 vs. 82 ms). La consonne pivot  $C_p$  et la voyelle post-consonantique ( $V_2$ ) sont allongées lorsqu'elles sont à l'initiale du mot, mais ces deux phénomènes n'apparaissent que pour la première mention ( $C_p$ :  $F[1, 31] = 6.98$ , 78 vs. 73 ms;  $V_2$ :  $F[1, 31] = 8.02$ , 71 vs 75 ms). Enfin, les éléments de la syllabe finale  $C_3$  et  $V_3$  ne sont aucunement affectés par la segmentation induite. Le seul résultat significatif est une différence selon l'exemplaire pour la consonne  $C_3$  ( $F[1, 31] = 4.45$ ).

Bien que ténues, les différences apparaissent systématiques et non liées à une classe particulière de segments (Figure 1). Plus de la moitié des observations pour  $C_p$  se situe sous la diagonale, quelle que soit la consonne, indiquant un allongement à l'initiale de mot (la condition de segmentation CCVC est en ordonnée). Inversement, 2/3 des observations pour la voyelle se

situent au-dessus de la diagonale, ce qui correspond également à une durée augmentée lorsque la voyelle est en début de mot.



**Figure 1:** Diagrammes de dispersion pour les durées de  $C_p$  (haut) et  $V_2$  (bas). En abscisse, les durées pour CCV#CCVCV, en ordonnée, pour CCVC#VCV. Les données correspondent à la moyenne des deux répétitions pour la première mention.

#### 3.2. Fréquence fondamentale

La table 3 indique les valeurs moyennes de  $F_0$  sur chaque syllabe, pour les deux mentions, en fonction de l'exemplaire et de la segmentation lexicale. On constate que la fréquence fondamentale varie entre les deux mentions ( $F[1, 31] = 1447.8$ ) et des patrons différents apparaissent (interaction Syllabe x Mention:  $F[1, 31] = 374.4$ ). Le contour mélodique montant de la première mention correspondrait à la mélodie typique de la question. Celui de la seconde mention, plus grave et légèrement descendant, indiquerait la fin de l'énoncé et pourrait constituer une marque de la familiarité de l'information sémantique portée. En revanche, aucune différence systématique n'apparaît, ni entre répétitions, ni

<sup>1</sup> Tous les résultats statistiques cités, sauf mention contraire, sont significatifs au seuil  $\alpha = .05$

en fonction de la segmentation lexicale induite.

**Table 3** : Fréquences fondamentales moyennes.

	Mention 1			Mention 2		
	V <sub>1</sub>	V <sub>2</sub>	V <sub>3</sub>	V <sub>1</sub>	V <sub>2</sub>	V <sub>3</sub>
CCV-R <sub>1</sub>	233.3	209.8	274.4	184.2	177.9	171.1
CCV-R <sub>2</sub>	229.0	210.1	275.5	184.8	177.4	171.4
CCVC-R <sub>1</sub>	236.9	213.1	279.9	184.4	175.6	171.0
CCVC-R <sub>2</sub>	228.8	207.8	273.4	183.9	176.6	172.3
Moyennes	232.0	210.2	275.8	184.3	176.9	171.4

### 3.3. Analyse spectrale

L'analyse formantique pratiquée pour chacune des voyelles de la séquence n'indique pas de variations majeures liées à la condition de frontière. Toutefois, il apparaît que le premier formant (F<sub>1</sub>) de la voyelle V<sub>1</sub> est plus élevé dans la condition CCVC#VCV que dans la condition CCV#CVCV (F[1, 31] = 7.47). Une légère tendance apparaît pour F<sub>2</sub>, mais elle n'est pas statistiquement significative (F[1, 31] = 2.05, p = .16). Aucun autre effet n'apparaît, hormis quelques différences entre mentions ou entre répétitions, que nous ne détaillerons pas. En résumé, seules les voyelles constituant le noyau de syllabes finales présentent des modifications en termes d'augmentation du degré d'aperture dans la condition d'enchaînement.

**Table 4** : Valeurs centrales des formants.

	Mention 1		Mention 2		
	CCV	CCVC	CCV	CCVC	
F <sub>1</sub>	V <sub>1</sub>	511	530	504	516
	V <sub>2</sub>	479	480	463	458
	V <sub>3</sub>	503	476	471	462
F <sub>2</sub>	V <sub>1</sub>	1628	1662	1670	1713
	V <sub>2</sub>	1864	1847	1824	1846
	V <sub>3</sub>	1617	1547	1623	1561
F <sub>3</sub>	V <sub>1</sub>	2752	2829	2755	2779
	V <sub>2</sub>	2898	2921	2829	2868
	V <sub>3</sub>	2796	2754	2812	2794

## 4. DISCUSSION

Ces résultats confirment que la réalisation acoustique des segments phonétiques varie selon la segmentation lexicale. En effet, tant la consonne pivot que la voyelle est allongée à l'initiale d'un mot. De plus, la voyelle pré-consonantique de la suite V<sub>1</sub>C<sub>p</sub>V<sub>2</sub> est temporellement et spectralement modifiée par la nature de la frontière suivante. Pour l'essentiel, nos résultats sont similaires à ceux rapportés par Fougeron et al. [6, 7]. Ils renforcent donc l'hypothèse que des indices phonétiques fins permettraient au locuteur de distinguer consonne initiale de mot et consonne finale enchaînée.

Les effets observés pour la consonne pivot et la voyelle V<sub>2</sub> ne semblent pas propres à des classes phonétiques particulières. Par contre, dans la mesure où la répartition de V<sub>1</sub> n'a pas été strictement contrôlée, les effets pour V<sub>1</sub> pourraient être liés à d'autres phénomènes, comme la loi

de position, par exemple, en interaction avec les intentions de segmentation des locuteurs.

Par ailleurs, l'importance des indices lexicaux semble varier selon les locuteurs [6, 7]. Une extension avec plusieurs locuteurs est donc indispensable afin de confirmer nos conclusions. De plus, des tests perceptifs seraient capitaux pour s'assurer que, comme dans le cas de la liaison (Spinelli et al., [8]) les indices présents sont effectivement utilisés dans la reconnaissance des mots.

En tout état de cause, cette première étude exploratoire valide l'usage de pseudo-mots dans l'étude des phénomènes phonétiques et phonologiques aux frontières de mots. D'une part, un tel corpus permet la mise en évidence de résultats étroitement similaires à ceux observés sur des séquences de mots [6, 7, 8]. D'autre part, il n'entrave pas le marquage linguistique; par exemple, les indices mélodiques typiques de la modalité linguistique (assertion vs. question) sont présents.

## REMERCIEMENTS

Cette recherche a été financée par le Fonds national suisse de la recherche scientifique (projet n° 1114-059532/5) et par le Fonds National de la Recherche Scientifique de Belgique (Convention FRFC n° 2.4610.04). Les auteurs remercient Julie Franck pour sa participation aux enregistrements.

## BIBLIOGRAPHIE

- [1] A. Content, R.K. Kearns, & U.H. Frauenfelder. Boundaries versus onsets in syllabic segmentation, *Journal of Memory and Language*, 45, 177-199, 2001.
- [2] A. Content & U.H. Frauenfelder. La syllabe comme unité de perception de la parole: un état de la question. JEP2002, Nancy, 24-27 June 2002.
- [3] R.A. Krakow. The Articulatory Organization of Syllables: A Kinematic Analysis of Labial and Velic Gestures. Ph.D. dissertation, Yale Univ., 1989.
- [4] H. Quené. Durational cues for word segmentation in Dutch. *Journal of Phonetics*, 20, 331-350, 1992.
- [5] N. Dumay, A. Content & U.H. Frauenfelder. Acoustic-phonetic cues to word boundary location: Evidence from word spotting. *Proceedings of the 14th ICPHS*, 281-284, San Francisco, 1999.
- [6] C. Fougeron, O. Bagou, M. Stefanuto, U.H. Frauenfelder. A la recherche d'indices de frontière lexicale dans la resyllabation, JEP2002, Nancy, 24-27 June 2002.
- [7] C. Fougeron, O. Bagou, A. Content, M. Stefanuto, U.H. Frauenfelder. Looking for acoustic cues of resyllabification in French, *Proceedings of the 15th ICPHS*, 2257-2260, Barcelone, August 2003.
- [8] E. Spinelli, J. McQueen, & A. Cutler. Processing resyllabified words in French. *Journal of Memory and Language*, 48, 233-254, 2003.