



Chapitre d'actes

1996

Published version

Open Access

This is the published version of the publication, made available in accordance with the publisher's policy.

La segmentation lexicale : contribution des indices segmentaux et suprasegmentaux

Meunier, Christine; Frauenfelder, Ulrich Hans; Content, Alain

How to cite

MEUNIER, Christine, FRAUENFELDER, Ulrich Hans, CONTENT, Alain. La segmentation lexicale : contribution des indices segmentaux et suprasegmentaux. In: XXIèmes Journées d'Etude sur la Parole. Avignon (France). [s.l.] : [s.n.], 1996. p. 31–34.

This publication URL: <https://archive-ouverte.unige.ch/unige:83176>

LA SEGMENTATION LEXICALE: CONTRIBUTION DES INDICES SEGMENTAUX ET SUPRASEGMENTAUX

Christine Meunier¹, Uli Frauenfelder¹ et Alain Content^{1,2}

¹ Laboratoire de Psycholinguistique Expérimentale,
FAPSE - Université de Genève - 9 route de Drize - CH-1227 Carouge - SUISSE
tél: (41) 22 705 97 41 - e-mail: meunier@fapse.unige.ch

² Laboratoire de Psychologie Expérimentale, Université libre de Bruxelles

ABSTRACT

Previous studies have shown the importance of prosodic features for lexical parsing. More precisely, the importance of rhythm was demonstrated in the segmentation of ambiguous bisyllabic words. This paper evaluates the weight of different features (duration, intonation, coarticulation, intensity) in lexical parsing. Our results show that temporal modifications, as well as other features, induce changes in word segmentation.

1. INTRODUCTION

Un des problèmes majeurs de la psycholinguistique est la compréhension des mécanismes qui entrent en jeu dans la reconnaissance des mots de la chaîne parlée. Si le code orthographique permet aisément d'isoler les mots dans une phrase, le signal acoustique ne laisse pas apparaître d'indices aussi évidents que les "blancs" séparant les mots sur une page de texte. Ainsi, le problème de la reconnaissance et donc l'extraction d'unités dans un flux continu nous renvoie directement au problème de la segmentation.

L'analyse acoustique permet de mettre en évidence plusieurs types d'indices relevant de niveaux d'analyse différents: d'une part les indices segmentaux (l'information allophonique et la coarticulation) et d'autre part, les indices suprasegmentaux ou prosodiques (durée, intonation et amplitude du signal).

L'utilisation des indices prosodiques par les auditeurs a clairement été démontrée pour l'anglais (Cutler et Norris, 1988). En français, Banel et Bacri (1994) ont montré que la segmentation lexicale est influencée par les indices rythmiques. Ces auteurs ont utilisé des mots bisyllabiques composés de deux monosyllabes enchassés. Les sujets devaient décider si l'expression présentée correspondait

à deux mots ou à un seul. En manipulant les durées relatives des deux voyelles, Banel et Bacri ont constaté que le rythme iambique (bref-long) favorisait le groupement des deux syllabes (réponse 1 mot) tandis que le rythme trochaïque (long-bref) augmentait leur séparation. Cependant, dans cette étude, les stimuli étaient toujours réalisés à partir de la concaténation des deux monosyllabes. On peut donc penser que les indices acoustiques non manipulés favori-saient la séparation. Toutefois, d'autres études ont montré l'effet de la durée en tenant compte des autres facteurs (Rietveld, 1980).

Notre objectif est d'évaluer la contribution relative des différents types d'indices (rythme d'une part, coarticulation, intonation et amplitude, d'autre part). Pour ce faire, nous avons modifié l'organisation temporelle de mots bisyllabiques, en partant soit du mot bisyllabique, soit d'une concaténation de leur composants monosyllabiques. Nous obtenons ainsi des stimuli pour lesquels les indices de durée entrent en conflit avec l'information induite par les autres indices (coarticulation, intonation et amplitude du signal). Deux expériences ont été menées, en variant la tâche proposée aux sujets.

2. PLAN EXPERIMENTAL

Quatre versions des stimuli ont été préparées afin de mettre en compétition différents types d'indices.

La première version (V1) est constituée par l'enregistrement naturel du mot bisyllabique (exemple: "corsage"). La deuxième version (V2) est constituée des deux mots monosyllabiques concaténés ("corps", "sage"). Les versions 3 et 4 sont obtenues en appliquant le schéma rythmique de V1 sur V2 et inversement (table 1):

Table 1: définition et élaboration des versions selon leur schéma rythmique et leur enchaînement segmental: en gras les versions "d'origine", en normal, les versions "modifiées".

	Bref-Long	Long-Long
non concaténés	V1 -----> V4	
concaténés	V3 <----- V2	

Ainsi, pour les versions 1 et 2 les indices phonétiques dont dispose l'auditeur sont congruents (table 2). A l'inverse, pour les versions 3 et 4, l'information temporelle est en conflit avec l'information apportée par les autres indices phonétiques. La compétition entre les différents types d'indices devrait donc entraîner une certaine ambiguïté dans le traitement de ces versions.

Table 2: type d'information (en terme de segmentation lexicale: 1 mot ou 2 mots) apportée par les différents indices pour les 4 versions

	V1	V2	V3	V4
durée	1	2	1	2
coarticulation	1	2	2	1
intonation	1	2	2	1
amplitude	1	2	2	1

La comparaison entre ces quatre versions nous permet de dégager la prédominance éventuelle de certains types d'indices pour la segmentation lexicale. Deux hypothèses s'offrent à nous:

1/ Il existe une dominance d'un type d'indice sur les autres et dans ce cas, soit: a) les indices temporels sont prédominants. V3 serait alors interprétée majoritairement comme 1 mot bisyllabique et V4 comme 2 mots monosyllabiques; b) les autres indices sont prédominants, et V3 serait interprétée comme 2 mots et V4 comme 1 mot.

2/ il n'y a pas de dominance entre les indices et dans ce cas l'interprétation de V3 et V4 est ambiguë.

3. EXPERIENCE I

3.1. Matériel linguistique

Nous avons sélectionné 24 mots bisyllabiques de la langue française dont chaque syllabe isolée forme un mot

monosyllabique (ex. "bulgare" / "bulle" "gare").

Les 24 mots bisyllabiques ainsi que les 48 mots monosyllabiques ont été enregistrés par un locuteur. La durée de chaque syllabe des mots monosyllabiques et des bisyllabiques a ensuite été mesurée. Ces mesures mettent en évidence une réduction notable de la durée de la première syllabe des bisyllabes (table 3). Ces observations sont conformes aux remarques de Rossi et al. (1981) qui montraient un fort allongement de la syllabe finale par rapport aux syllabes précédentes, à l'intérieur d'un groupe rythmique.

Table 3: durées moyennes (en ms) des syllabes S1 et S2 dans les mots bisyllabiques et monosyllabiques.

	bisyll. (ex. bulgare)	monosyll. (ex. bulle gare)
S1	233	453
S2	366	359

Les versions 1 et 2 constituent les versions "naturelles". Les versions 3 et 4 sont obtenues à l'aide d'un calcul mettant en évidence les relations temporelles liant S1 et S2 dans les couples de monosyllabes et dans les mots bisyllabiques. Ce sont les versions "modifiées".

Version 3

La version 3 est obtenue en appliquant le schéma temporel du mot bisyllabique ("bulgare") sur la suite de monosyllabes ("bulle"+"gare"): la durée du mot monosyllabique correspondant à S1 est réduite de façon à ce que le rapport S1/S2 soit comparable à celui que l'on trouve dans le mot bisyllabique.

Cette formule est appliquée à chaque item de façon à ce que le rapport S1/S2 soit spécifique à chaque mot. Cette précaution nous semble nécessaire dans la mesure où les rapports S1/S2 sont assez variables d'un item à l'autre en fonction des unités phonétiques présentes.

Version 4

La relation inverse est appliquée de façon à obtenir la version 4 à partir d'une transformation de V1. Dans V4, le mot "bulgare" reçoit ainsi un schéma temporel

comparable à celui de "bulle"+"gare" dans V2, par allongement de la 1ère syllabe.

L'allongement et la compression ont été réalisés à l'aide du logiciel SoundDesigner II.

3.2. Procédure

Les stimuli ont été enregistrés sur cassette. Chaque stimulus était présenté 3 fois avec un intervalle d'une seconde.

L'ensemble des 96 stimuli ont été répartis en 4 listes, chacune contenant l'ensemble des 24 stimuli. Les 4 versions de chaque item ont été réparties de manière équilibrée entre les listes.

Les sujets (n=49) ont été répartis en 4 groupes. Chaque groupe a entendu une des 4 listes, précédée de 4 exemples.

Les sujets devaient transcrire ce qu'ils entendaient et avaient la possibilité de proposer plusieurs transcriptions en les classant par ordre de préférence. Le temps de réponse des sujets était limité à 10 secondes.

3.3. Résultats

Les résultats de cette première expérience montrent que pour les deux versions non modifiées (V1 et V2) les réponses des sujets expriment clairement l'information apportée par les deux types d'indices (table 2): V1 est interprétée comme 1 mot bisyllabique et V2 majoritairement comme 2 mots monosyllabiques (figure 1).

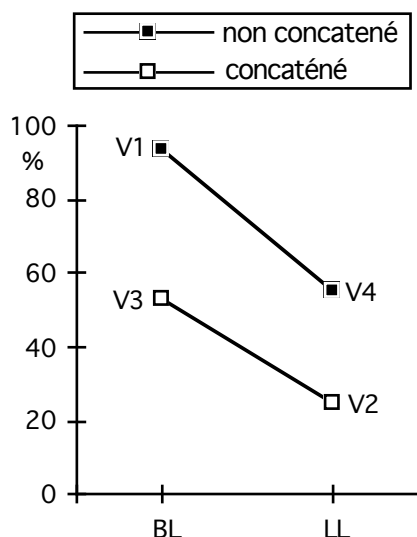


Figure 1: Taux de réponses "un mot".

Pour les deux versions modifiées (V3 et V4), les résultats sont moins contrastés: pour

chacune des versions on observe autour de 50% de réponses "1 mot".

La modification du rythme dans un mot bisyllabique (V1->V4) réduit de 40% le taux de groupement ($p<.0001$), tandis que la modification rythmique inverse dans deux mots monosyllabiques (V2->V3) réduit de 30% le taux de séparation ($p<.0001$). Par ailleurs, la modification des autres indices (coarticulation, intonation, intensité) entraîne également des différences significatives—pour une même structure rythmique (V1->V3, $p<.0001$; V2->V4, $p<.0001$).

La tâche de choix libre donne lieu à 6.8% de réponses "non attendues", c'est-à-dire celles qui ne correspondent ni au groupement ni à la séparation (V1: 3.5%, V2: 4.6%, V3: 9.6%, V4: 9.6%). On observe donc, dans l'ensemble, que la majeure partie des réponses se répartissent entre les deux catégories attendues.

Au vu de cette première expérience, il semble que la deuxième hypothèse exposée plus haut soit observée: il n'apparaît pas de dominance des différents types d'indices, puisque la manipulation du rythme montre un effet aussi bien pour la version concaténée que pour la version non-concaténée.

Une deuxième expérience a été mise au point afin d'approfondir et de préciser les résultats de l'expérience 1. La tâche de choix forcé devrait éliminer les réponses non attendues et donc augmenter le pourcentage de réponses 1 mot pour V1 et de réponses 2 mot pour V4. Par ailleurs, la mesure des temps de réaction (TR) devrait nous permettre de confirmer la difficulté de traitement des stimuli ambigus.

4. EXPÉRIENCE 2

4.1. Matériel linguistique

Les stimuli de l'expérience 2 sont identiques à ceux de l'expérience 1.

4.2. Procédure

L'ensemble des 25 sujets ont entendu la totalité des stimuli. Les 4 listes, et les items dans chaque liste, étaient présentées dans un ordre aléatoire.

Les sujets avaient une présentation visuelle des deux possibilités (1 mot bisyllabique ou 2 mots monosyllabiques) et devaient, après audition du stimulus, choisir l'une des alternatives en appuyant sur l'un des deux boutons d'un boîtier de réponse le plus rapidement possible. L'expérience a été réalisée à l'aide du logiciel Psyscope.

4.3. Résultats

Les résultats confirment et renforcent ceux obtenus pour l'expérience 1. Cette fois, les sujets sont unanimes à interpréter la version "mot bisyllabique" comme 1 seul mot, et la version "mots concaténés" comme une suite de deux mots monosyllabiques (figure 2). La tâche de choix forcé a donc contrasté les réponses concernant les deux versions naturelles.

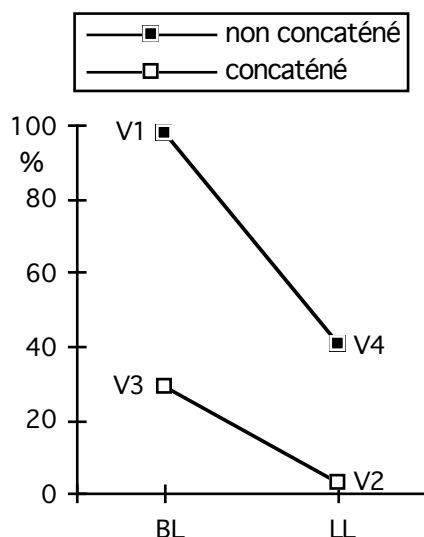


Figure 2: Taux de réponse "un mot"

Les résultats pour les deux manipulations du rythme donnent lieu à des taux de réponses intermédiaires. Les deux manipulations opposées ont des effets systématiques et statistiquement significatifs. L'allongement de la première syllabe dans les mots bisyllabiques réduit de plus de 50% le taux d'interprétations "mot bisyllabique" ($p < .0001$). Inversement, un rythme bref-long sur les paires de mots concaténés augmente de 25% l'interprétations "mot bisyllabique" ($p < .0001$). Les variations de réponses dues aux facteurs communs (coarticulation, intonation, intensité) sont également significatives. Enfin, les résultats montrent une interaction ($p < .0002$), puisque l'effet de la manipulation du rythme est plus

marqué sur la version non-concaténée, qui préserve les indices de coarticulation, d'intonation et d'amplitude. Ce résultat confirme donc l'importance de l'indice de durée pour la segmentation lexicale.

Toutefois, il faut noter une très grande variabilité des réponses pour les versions modifiées. Cette variabilité semble être fonction des stimuli.

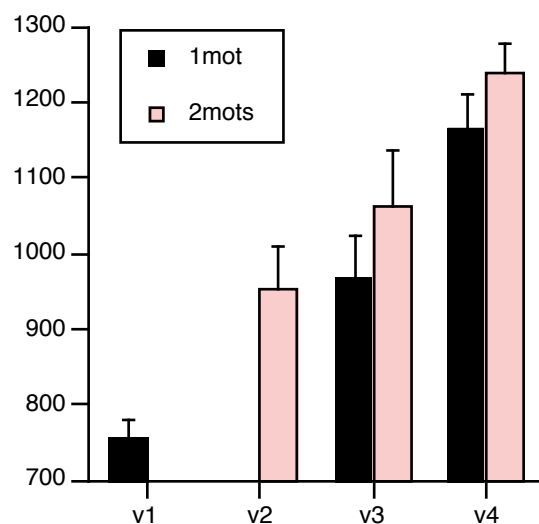


Figure 3: temps de réaction moyen (en ms) des réponses "1 mot" et "2 mots".

Les TR observés sont systématiquement fonction du degré d'ambiguïté du stimulus. Les temps les plus courts sont observés dans les deux conditions où les sujets montrent la plus grande unanimité (figure 3, V1: 755ms; V2: 953ms). Entre les deux conditions intermédiaires, les temps sont plus longs dans la situation bisyllabes long-long (V4), qui correspond au taux de réponses le plus proche de 50%.

Ces observations nous permettent de confirmer la difficulté de traitement des stimuli ambigus.

5. DISCUSSION ET CONCLUSION

Les conditions énumérées au début de l'article permettaient d'envisager deux hypothèses: soit les informations apportées par les différents indices étaient traitées différemment selon le poids des indices, soit elles ne l'étaient pas.

L'hypothèse d'indices dominants n'a pas été confirmée. Il semble effectivement que les informations apportées par les indices

temporels, d'une part, et celles apportées par les autres indices, d'autre part, ne permettent pas un type de réponse majoritaire lorsqu'elles sont en contradiction.

Ainsi, lorsque toutes les informations sont congruentes, le traitement est simple, lorsqu'elles sont en conflit, l'interprétation du signal devient difficile et contradictoire.

Il nous faut toutefois nuancer ces propos en évoquant la très forte variabilité des réponses en fonction des items. Il semble possible que la conjonction de certains facteurs favorise le poids d'un indice. Par ailleurs, une manipulation plus approfondie et plus fine des stimuli pourrait sans doute nous permettre de mieux comprendre cette forte variabilité. En particulier, il pourrait être intéressant de tenter de dissocier expérimentalement les effets des indices d'intonation, d'amplitude et de coarticulation que nous n'avons pas distingués dans cette première exploration.

L'intérêt de ces expériences est de mettre en opposition différents types d'indices phonétiques afin d'évaluer leur contribution pour la segmentation lexicale. Nous pensons par la suite poursuivre ces expériences en utilisant des tâches moins directement liées à la segmentation (comme une tâche de décision lexicale, par exemple).

6. BIBLIOGRAPHIE

Banel M.H., Bacri N. (1994) "On metrical patterns and lexical parsing in French", *Speech Communication*, 15, pp. 115-126.

Cutler A., Norris D. (1988) "The role of strong syllables in segmentation for lexical access", *J. Experimental Psychology: Human Perception and Performance*, Vol. 14, pp. 113-121.

Rietveld A.C.M. (1980) "Word boundaries in the French language", *Language and Speech*, Vol. 23, No. 3, pp. 289-296.

Rossi M., Di Cristo A., Hirst D.J., Martin P., Nishinuma Y. (1981) *L'intonation: de l'acoustique à la sémantique*. Paris, Klincksieck.

Remerciements

Cette recherche a été financée par le F.N.R.S. 11-39553.93. Nous remercions Marie Bachmann et Isabelle Gunther pour leur aide à la réalisation de l'expérience 1.