

Influence du rythme sur le traitement sémantique en français: Approches comportementale et électrophysiologique

Corine Astésano, Cyrille Magne, Radouane El Yagoubi & Mireille Besson

Institut de Neurosciences Cognitives de la Méditerranée, CNRS, 31 Ch. J. Aiguier, 13402 Marseille Cedex 20, France

Tél.: ++33 (0)4 91 16 43 94 - Fax: ++33 (0)4 91 77 49 69

Mél: {astesano; cmagne; mbesson}@lnf.cnrs-mrs.fr ; Radouane.El-Yagoubi@up.univ-mrs.fr

ABSTRACT

The present study aims at studying the relationship between rhythmic (lexical meter) and semantic processing under two attention tasks (semantic and rhythmic), using the Event-Related Potential (ERP) method. Participants listened to short sentences ended by trisyllabic semantically and/or rhythmically congruous or incongruous words. Preliminary results indicate that, in the rhythmic task, the metric structure is processed on-line as reflected by the occurrence of an increased positivity to rhythmic incongruities between 500 and 800 ms. Moreover, results in the semantic task reveal that an N400 component is associated to semantically congruous but rhythmically incongruous final words. Thus, this latter result is indicative of the influence of the metric structure on lexical access in French.

1. INTRODUCTION

La prosodie joue un rôle central d'organisation et de structuration du message dans le langage parlé. Ces fonctions sont réalisées à travers les patrons accentuels et intonatifs qui délimitent non seulement le niveau du mot, par le marquage métrique du lexique, mais aussi le niveau des groupements de mots, par le marquage de la structure syntaxique ('phrasing'). L'accentuation structure également l'énoncé par la mise en valeur acoustique des unités de sens les plus importantes (focalisation de mots ou de parties du discours). La prosodie intervient donc à tous les niveaux d'organisation du langage: lexique et sémantique, structure syntaxique et structure informationnelle (pragmatique). Cependant, bien que la prosodie soit une caractéristique fondamentale du langage humain, sa contribution à la compréhension en temps réel du langage parlé a reçu relativement peu d'attention dans les modèles cognitifs, par rapport à d'autres aspects linguistiques, tels que la syntaxe ou la sémantique (voir Besson et al. [4] pour une revue de la littérature). Cette absence s'explique en partie par la complexité intrinsèque de la prosodie. En effet, les mêmes indices acoustiques (F0, durée, intensité) sont utilisés conjointement pour transmettre les différentes fonctions de la prosodie (lexicales, syntaxiques, pragmatiques). Il reste donc très difficile de séparer le

traitement de la prosodie *per se* des autres niveaux linguistiques avec lesquels elle interagit.

Au cours des cinq dernières années, plusieurs études ont cependant eu pour but d'étudier, en temps réel, le déroulement temporel du traitement prosodique en relation avec les traitements syntaxique ou sémantique.

En utilisant la méthode des Potentiels Evoqués (PEs), qui permet une mesure directe de l'activité électrique cérébrale associée à la présentation d'un événement particulier [3] et qui est connue pour son excellente résolution temporelle, il a ainsi pu être démontré qu'une composante positive se développe au niveau des frontières prosodiques, en relation avec le traitement syntaxique (e.g. [11]). En outre, dans une étude où incongruïtés sémantique et prosodique (contours mélodiques dans la fonction de marquage de la modalité des phrases) ont été manipulées orthogonalement, Astésano, Besson & Alter [2] ont montré que le traitement de l'information de modalité, associé à l'occurrence d'une composante positive tardive (P800), est modulé par le traitement de la sémantique (composante N400). Ainsi, lorsque les sujets focalisent leur attention sur les propriétés mélodiques de la phrase, l'effet d'incongruïté prosodique est plus fort (composante P800 plus ample) lorsque la sémantique est également incongrue. En revanche, l'incongruïté prosodique ne module pas le traitement de la sémantique. Ainsi, la sémantique serait traitée de manière obligatoire, ce qui ne serait pas le cas des informations prosodiques.

Le but de la présente expérience est à nouveau d'examiner les relations entre traitements sémantique et prosodique mais d'analyser cette fois le rôle de la prosodie dans sa fonction de marquage métrique du lexique. Notons en effet, que des descriptions récentes du français proposent que le mot est marqué par un Accent Initial (mélodique) et un Accent Final (allongement, [6 ; 1]) Cependant, l'accent initial étant secondaire, le patron métrique de base serait iambique, l'allongement final seul marquant le mot ou le groupe de mots. Afin de manipuler ce patron métrique de base et de créer des incongruïtés rythmiques, l'allongement syllabique final a été déplacé à la pénultième syllabe de mots trisyllabiques situés à la fin de phrases sémantiquement congruentes ou incongrues. Il est ainsi possible de comparer le déroulement temporel et la distribution topographique des effets de congruence

prosodique et sémantique. Enfin, en manipulant l'attention des participants (tâche de jugement sémantique ou tâche de jugement rythmique, du dernier mot des phrases), cette étude nous permet d'analyser les effets d'un traitement implicite ou explicite du marquage métrique [2]. Dans cette optique, nous avons enregistré conjointement les indices comportementaux tels que les temps de réaction (TRs) et les taux d'erreurs (%Err) et les indices électrophysiologiques (PEs). Dans cet article court, la présentation et la discussion des résultats préliminaires sera essentiellement centrée sur le traitement explicite (tâche rythmique) ou implicite (tâche sémantique) du rythme.

2. MATÉRIEL ET MÉTHODE

2.1. Matériel linguistique

Un corpus de 128 phrases, terminées par un mot trisyllabique (e.g. *Le concours a regroupé mille candidats*) a été construit pour cette expérience. Les phrases ont été prononcées à haute voix par un locuteur masculin, et enregistrées en chambre anéchoïque sur un support DAT (i.e., Digital Audio Tape) à une fréquence d'échantillonnage de 44,1 kHz.

Congruences rythmique et sémantique ont été manipulées de façon orthogonale, afin de construire quatre conditions expérimentales dans lesquelles le dernier mot trisyllabique des phrases était : 1) rythmiquement et sémantiquement congruent (R+S+), 2) rythmiquement congruent et sémantiquement incongru (R+S-), 3) rythmiquement incongru et sémantiquement congruent (R-S+), 4) rythmiquement et sémantiquement incongru (R-S-). Chaque condition expérimentale comprenait 32 phrases (cf. Table 1).

Table 1: Exemple de phrase-test (R+ = rythme congruent; S+ = sémantique congruente; R- = rythme incongru; S- = sémantique incongrue). Le soulignement dans la colonne R- indique les syllabes artificiellement allongées pour créer l'incongruité rythmique.

	R+	R-
S+	Le concours regroupait mille candidats	Le concours regroupait mille candidats
S-	Le concours regroupait mille bigoudis	Le concours regroupait mille bigoudis

L'incongruité sémantique (S-) consistait à remplacer le dernier mot des phrases congruentes par un mot qui n'était pas attendu dans le contexte de la phrase. L'incongruité rythmique (R-) consistait à allonger la pénultième syllabe du dernier mot de la phrase d'une valeur de 1.9, correspondant à la moyenne d'allongement de la dernière syllabe des mots des groupes prosodiques en français [1]. L'implémentation

de l'allongement s'est faite à l'aide d'un algorithme permettant de manipuler la durée d'un signal acoustique, sans en modifier le timbre ou la fréquence [10].

2.2. Participants

Sept adultes droitiers, d'âge moyen 26 ans, ont été rémunérés pour participer à cette expérience. Tous les participants avaient une audition correcte et étaient de langue maternelle française.

2.3. Procédure et déroulement de l'expérience

Dans la tâche sémantique, les participants devaient focaliser leur attention sur le sens de la phrase afin de déterminer si le dernier mot était congruent ou incongru dans le contexte. Dans la tâche rythmique, les participants devaient focaliser leur attention sur le rythme du dernier mot, afin de déterminer si le rythme syllabique était correct ou non. L'ordre des tâches était contre-balancé à travers les participants afin d'éviter tout effet d'ordre. Les sessions expérimentales comprenaient 4 blocs d'essais (2 blocs par type de tâche, 32 phrases par bloc) et étaient précédées de 10 phrases de familiarisation avec les tâches. Dès la présentation du dernier mot de la phrase, le participant devait donner sa réponse, le plus rapidement possible et en faisant le moins d'erreurs possibles, en appuyant sur un bouton de réponse.

2.4. Enregistrement et analyse des données

Les PEs ont été enregistrés à partir de 28 électrodes intégrées dans un casque (ElectroCap) et placées sur les sites typiques du système 10/20 de la Fédération Internationale d'électroencéphalographie [8]. La référence commune est placée sur la mastoïde gauche et une électrode est également placée sur la mastoïde droite afin de moyennner, off-line, les effets enregistrés par rapport à la moyenne de l'activité électrique sur les deux mastoïdes. Les mouvements horizontaux des yeux et les saccades sont enregistrés grâce à deux électrodes situées sur les canthi externes et une électrode placée sous l'œil gauche. L'enregistrement de l'électroencéphalogramme (EEG) commence 200 ms avant la présentation du dernier mot et continue pendant 2200 ms, avec une fréquence d'échantillonnage de 250 Hz. Les essais sont ensuite moyennnés afin d'obtenir les PEs, après que ceux comportant des mouvements oculaires aient été rejetés.

Les différences observées entre chaque condition expérimentale (i.e. conditions R+S-, R-S+, et R-S-) et la condition de complète congruence (i.e. R+S+) ont été testées statistiquement 2 à 2, à l'aide d'analyses de variance (ANOVAs). Les analyses statistiques ont été effectuées sur l'amplitude moyenne des PEs enregistrés sur les électrodes médianes du scalp (i.e. Fz, Cz, Pz, Oz) et dans des rangs de latence déterminés

d'après inspection visuelle des données. Les facteurs conditions (2) et électrodes (4) ont été utilisés pour les ANOVAs.

3. RÉSULTATS

3.1. Données comportementales

Les résultats n'ont révélé aucune différence significative, ni sur les temps de réaction, ni sur les pourcentages d'erreur ($F < 1$). Les résultats comportementaux semblent cependant indiquer que lorsque l'attention du participant est focalisée sur la sémantique, le TR est plus long pour décider que le dernier mot est sémantiquement congruent, lorsqu'il est également rythmiquement incongru (condition R-S+ : 1042 ms), que lorsqu'il est rythmiquement congruent (condition R+S+ : 926 ms).

3.2. Données électrophysiologiques

Lorsque l'attention du participant est focalisée sur le rythme, les mots sémantiquement congruents et rythmiquement incongrus (R-S+) suscitent l'occurrence d'une composante positive (P600) plus

ample que les mots rythmiquement et sémantiquement congruents (R+S+), entre 500 et 800 ms après le début du mot [effet principal du facteur condition : $F(1,6) = 9.18$; $p = 0.02$]. Cette composante est d'amplitude maximale sur les régions pariétales du scalp (interaction condition par électrode : $F(3,18) = 5.67$; $p = 0.017$; voir Figure 1, colonne gauche). Lorsque l'attention du participant est focalisée sur la sémantique, les mots sémantiquement congruents et rythmiquement incongrus (R-S+) suscitent l'occurrence d'une composante N400 plus ample que les mots rythmiquement et sémantiquement congruents (R+S+) entre 350 et 500 ms [effet principal du facteur condition : $F(1,6) = 26.55$; $p = 0.002$]. La composante N400 est d'amplitude maximale sur toutes les électrodes testées (région centrale du scalp ; voir Figure 1, colonne droite).

En revanche, dans les deux tâches, Rythme et Sémantique, les mots rythmiquement congruents mais sémantiquement incongrus (R+S-) suscitent l'occurrence d'une composante N400 plus grande que les mots rythmiquement et sémantiquement congruents (R+S+), entre 350 et 500 ms, avec un maximum d'amplitude sur les régions centrales du scalp

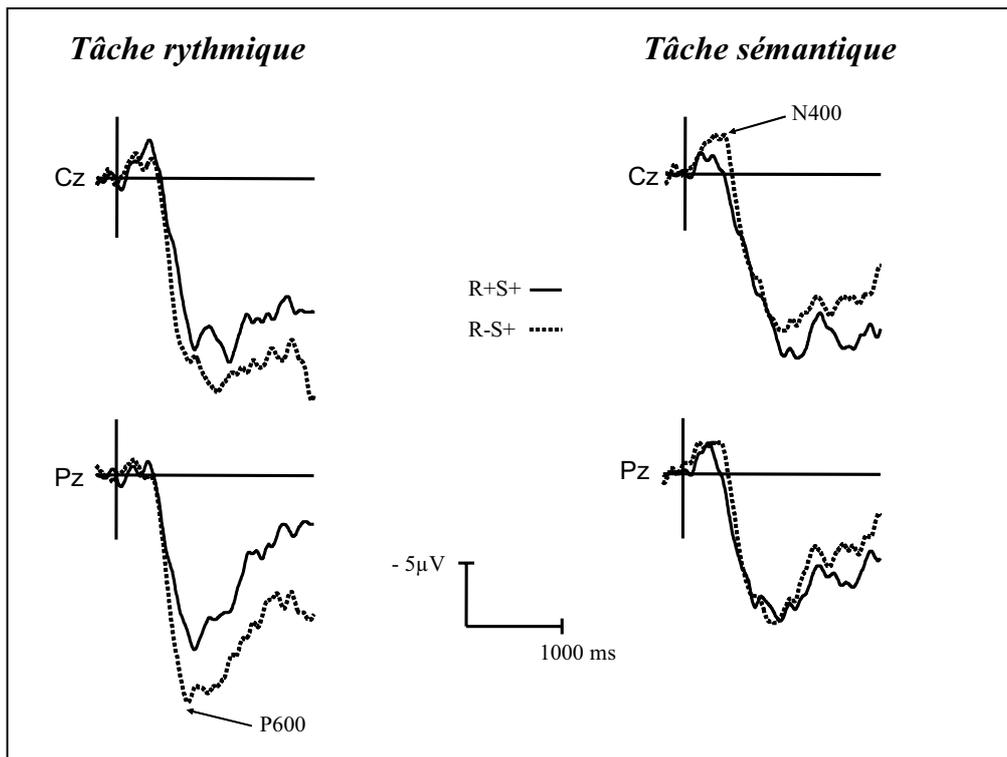


Figure 1: Grand-moyennage des données enregistrées chez 7 participants. Potentiels Evoqués (PEs) associés à la présentation du dernier mot sémantiquement congruent de la phrase, lorsqu'il est rythmiquement congruent (R+S+, tracé plein) ou rythmiquement incongru (R-S+, tracé en pointillé). La colonne de gauche présente les PEs lorsque le sujet réalise une tâche explicite sur le rythme, tandis que la colonne de droite présente les PEs lorsque le sujet réalise une tâche sémantique. Les données présentées sont enregistrées au niveau des électrodes centrale (Cz) et pariétale (Pz) de la ligne médiane du scalp. L'amplitude des phénomènes enregistrés est représentée en ordonnée (en microvolt, μV) et le temps en abscisse (en millisecondes, ms). La négativité est représentée vers le haut.

[Sémantique : effet principal du facteur condition : $F(1,6) = 13.65$; $p = 0.01$] ; Rythme : interaction condition par électrode : $F(3,18) = 3.89$; $p = 0.04$].

4. DISCUSSION

Cette étude des relations entre traitement de la structure métrique des mots et traitement sémantique permet de mettre en évidence plusieurs résultats intéressants. Premièrement, nous avons pu démontrer que, comme le sens, la structure métrique est traitée en temps réel. En effet, lorsque l'attention des participants est focalisée sur le rythme (tâche explicite), les mots rythmiquement incongrus suscitent l'occurrence d'une composante P600 plus ample que les mots rythmiquement congruents. Cette composante positive, suscitée entre 500 et 800 ms après le début du mot, soit environ 300-600 ms après le début de la deuxième syllabe, ressemble à la composante P300, associée à la présentation de stimuli peu fréquents ou qui provoquent un effet de surprise [7]. Ainsi, le fait qu'un allongement de la pénultième syllabe du dernier mot des phrases suscite un effet de surprise chez les participants suggère qu'ils perçoivent la déviation métrique et, par conséquent, qu'ils possèdent une représentation de la structure métrique du français. Deuxièmement, lorsque l'attention des participants est focalisée sur la sémantique et que les mots sont sémantiquement congruents, l'incongruité métrique suscite néanmoins l'occurrence d'une composante N400. Ainsi, si le mot comprend un allongement de la pénultième syllabe, cette incongruité métrique semble perturber l'accès au sens du mot, ce qui serait reflété par l'occurrence d'une composante N400. Ces résultats, qui devront être confirmés sur un plus grand nombre de sujets, semblent néanmoins indiquer que la structure métrique est traitée de manière implicite et influence l'accès au sens du mot. Ils soulignent donc l'importance de la fonction lexicale de la prosodie (marquage métrique).

Enfin, ces résultats montrent qu'indépendamment de la tâche effectuée par le participant (rythmique ou sémantique), les mots rythmiquement congruents mais sémantiquement incongrus suscitent l'occurrence d'une composante N400 [9]. Ainsi, les participants ne pourraient pas éviter de traiter le sens des mots même lorsque leur attention est centrée sur le rythme. En ce sens, ils confirment et étendent les résultats obtenus par Astésano et al. [2] soulignant le caractère obligatoire du traitement sémantique. Ils établissent également que le traitement explicite des aspects sémantiques et rythmiques est reflété par des manifestations électrophysiologiques différentes, respectivement les composantes N400 et les composantes positives tardives (P600 ou P800).

Concernant la suite de cette étude, il sera évidemment important de passer un plus grand nombre de participants afin de confirmer les résultats obtenus chez les 7 premiers. Il sera également intéressant de

déterminer dans quelle mesure la négativité suscitée par le traitement de la structure métrique reflète le traitement prosodique per se ou reflète plutôt la difficulté d'accès au sens liée à l'introduction de la déviation métrique du mot lexical. Afin de répondre plus en détail à cette question, des tâches de jugement implicite sont envisagées, comme le priming auditif de patrons métriques (à la suite de Böcker et al. sur le hollandais [5]) ou des tâches de jugement 'pareil/différent', dans lesquelles les sujets devraient décider si des mots identiques portant des patrons métriques différents sont les mêmes ou non.

BIBLIOGRAPHIE

- [1] C. Astésano. *Rythme et accentuation en français. Invariance et variabilité stylistique*. Éditions L'Harmattan, Paris, 2001.
- [2] C. Astésano; M. Besson & K. Alter. Brain potentials during semantic and prosodic processing in French. *Cognitive Brain Research*, 18(2):172-184, 2003.
- [3] H. Berger. Uber das Electrenkephalogram das menchen. *Archiv für Psych.*, 87:527-570, 1929.
- [4] M. Besson ; C. Magne & P. Regnault. Le traitement du langage. In B. Renault (Ed.), *L'imagerie fonctionnelle électrique (EEG) et magnétique (MEG): ses applications en sciences cognitives*. Editions Hermès, sous presse.
- [5] K.B.E. Böcker; M.C.M. Bastiaansen; J. Vroomen; C.H.M. Brunia & B. de Gelder. An ERP correlate of metrical stress in spoken word recognition. *Psychophysiol.*, 36:720-760, 1999.
- [6] A. Di Cristo. Le cadre accentuel du français contemporain: essai de modélisation (1^{ère} partie). *Langues*, 2 (3) :184-205, 1999.
- [7] E. Donchin. Surprise! ... Surprise? *Psychophysiol.*, 18:493-513, 1981.
- [8] H. Jasper. The 10-20 electrode system of the International Federation. *Electroencephalographie and Clinical Neurophysiology*, 10:370-375, 1958.
- [9] M. Kutas & S.A. Hillyard. Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207:203-205, 1980.
- [10] G. Pallone ; P. Boussard; L. Daudet ; P. Guillemain & R. Kronland-Martinet. A wavelet based method for audiovideo synchronization in broadcasting applications. *Proceedings of the DAFX'99*, Trondheim, Norway, 1999.
- [11] K. Steinhauer; K. Alter & A.D. Friederici. Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Neuroscience*, 2 (2) :191-196, 1999.