

Traitement de l'accent de focalisation en français: Approche électrophysiologique

Cyrille Magne¹, Corine Astesano¹, Anne Lacheret-Dujour², Michel Morel² & Mireille Besson¹

¹Institut de Neurosciences Cognitives de la Méditerranée, CNRS, 31 Ch. J. Aiguier, 13402 Marseille Cedex 20, France

Tél.: ++33 (0)4 91 16 41 13 - Fax: ++33 (0)4 91 77 49 69

Mél: {cmagne; astesano; mbesson}@inf.cnrs-mrs.fr

²CRISCO, Université de Caen, place de la Paix, 14032 Caen cedex, France

Tél.: ++33 (0)2 31 56 56 27 - Fax: ++33 (0)2 31 56 54 27

Mél: {anne.lacheret ; morel}@crisco.unicaen.fr - <http://www.crisco.unicaen.fr>

ABSTRACT

The present experiment was aimed at testing the psychobiological validity of prosodic focus in French, using the Event-Related Potential method (ERPs). The brain electrical activity of 16 subjects was recorded while they listened to short dialogues (one question and one answer). Four conditions were created, in which the answers included pragmatically congruous or incongruous focal accents, either on middle or final words. Results revealed that prosodic focus is processed online by the brain. Indeed, clear differences were found between pragmatically congruous and pragmatically incongruous focal accents. Moreover, the effects of pragmatic incongruities were modulated by the position of the focal accent in the sentence (P300 for middle words vs N400 for terminal words).

1. INTRODUCTION

Lors d'une situation de communication, telle qu'un dialogue, le locuteur organise son discours de telle sorte que l'information nouvelle qu'il veut communiquer, soit mise en exergue par rapport au reste du message. Ce processus par lequel une information nouvelle, perceptivement saillante, devient saillante du point de vue cognitif correspond au concept pragmatique de « focus ». Dans de nombreuses langues, notamment le français, le focus est exprimé prosodiquement via des paramètres acoustiques dans le signal de parole (durée, F0 et intensité).

Le focus prosodique est un aspect important du langage parlé, qui a été largement étudié par les linguistes et les phonéticiens, mais très peu abordé au moyen des méthodes d'imagerie cérébrale. Ainsi, certains auteurs soulignent que la plupart des résultats de la littérature reposent sur des études utilisant des méthodes de mesure indirecte (i.e., réponses comportementales), et qu'il serait important d'utiliser des méthodes de mesure directe afin de mieux cerner les mécanismes impliqués dans le traitement des indices prosodiques [1].

Le but de cette étude est de tester la validité psychobiologique du focus prosodique en français, en utilisant la méthode des Potentiels Evoqués (PEs). Cette méthode permet, à l'aide d'électrodes placées sur le scalp, une mesure directe de l'activité électrique cérébrale associée à la présentation d'un événement particulier [2]. Depuis le début des années 80, cependant, la plupart des études utilisant cette technique se sont focalisées sur les aspects sémantique [3] et syntaxique [4] du langage. Seules, quelques récentes études ont eu pour but d'étudier les corrélats électrophysiologiques de certains aspects prosodiques. Mais tandis que ces études se sont focalisées sur les fonctions lexicales [5], structurales [6], ou sémantique [7, 8], de la prosodie au niveau de la phrase, le but de cette étude est d'examiner le traitement du focus prosodique au niveau du discours, une situation plus écologique. En outre, nous avons travaillé sur un type particulier de focus prosodique, le focus contrastif. Il induit une sélection exclusive dans une classe paradigmatique [9]. Ainsi, la question « Est-ce que Jean a donné un bracelet ou une bague à sa fiancée ? » induit des attentes spécifiques sur les mots « bracelet » et « bague ». Ces attentes sont ensuite résolues grâce à la réalisation d'un accent focal sur le mot « bracelet » et non « bague » dans la réponse « Jean a donné un **bracelet** à sa fiancée ».

Nous avons créé des dialogues courts comprenant une question et une réponse comme dans l'exemple précédemment cité. De plus, pour chaque réponse, deux versions de la question ont été créées. Dans une version, l'accent focal contrastif était attendu sur le mot du milieu, alors que dans l'autre version, il était attendu sur le mot final de la réponse. Ainsi, la cohérence du dialogue dépendait du mot focalisé dans la réponse, en fonction des attentes générées par la question. En croisant la présentation des deux types de question (accent focal attendu sur le mot médian vs mot final), avec la présentation des deux types de réponse (accent focal sur le mot médian vs mot final), nous avons créés des situations dans lesquelles l'accent focal était mal placé soit au milieu, soit à la fin de la réponse. En résumé, nous avons utilisé 4 conditions

expérimentales (voir Table 1) dans lesquelles l'accent focal était :

- 1) bien placé au milieu et absent (non attendu) à la fin de la réponse (i.e., M+T0) ;
- 2) mal placé au milieu car attendu (mais absent) à la fin (i.e., M-T0) ;
- 3) absent au milieu (non attendu) et bien placé à la fin (i.e., M0T+) ;
- 4) mal placé à la fin car attendu (mais absent) au milieu (i.e., M0T-).

Table 1 : Exemple de dialogue utilisé dans chacune des 4 conditions expérimentales. Le mot focalisé dans la réponse est indiqué en italique.

Conditions	Exemples
M+T0	Qu'est ce que Jean a donné à sa fiancée, un bracelet ou une <i>bague</i> ? Jean a donné une <i>bague</i> à sa fiancée
M-T0	A qui Jean a-t-il donné une <i>bague</i> , à sa fiancée ou à sa soeur ? Jean a donné une <i>bague</i> à sa fiancée
M0T+	A qui Jean a-t-il donné une <i>bague</i> , à sa fiancée ou à sa soeur ? Jean a donné une <i>bague</i> à sa <i>fiancée</i>
M0T-	Qu'est ce que Jean a donné à sa fiancée, un bracelet ou une <i>bague</i> ? Jean a donné une <i>bague</i> à sa <i>fiancée</i>

2. MATÉRIEL ET MÉTHODE

2.1. Participants

16 adultes droitiers, d'âge moyen 27 ans, ont été rémunérés pour participer à cette expérience. Tous les participants ont une audition correcte et sont de langue maternelle française.

2.2. Stimuli

Un total de 240 questions et 240 réponses a été construit pour l'expérience (voir Astésano et al. dans ces actes [10] pour les détails sur les conditions d'enregistrement et l'analyse des paramètres acoustiques des dialogues). Pour une moitié des questions, l'accent focal contrastif était attendu sur le mot médian de la réponse, alors que pour l'autre moitié des questions, l'accent focal était attendu sur le mot final des réponses. Ainsi, chaque réponse a été enregistrée deux fois en fonctions de la place attendue de l'accent focal. En outre, 4 différentes listes de dialogues ont été créées de telle sorte que chaque participant entende chaque paire de question-réponse une seule fois, mais que toutes les paires question-réponse soit présentées dans les 4 conditions expérimentales à travers tous les participants. Chaque liste comprenait 120 dialogues : 60 étaient cohérents

(i.e., accent focal bien placé sur le mot médian ou le mot final) et 60 étaient incohérents (i.e., accent focal mal placé sur le mot médian ou le mot terminal).

2.3. Procédure

Durant l'expérience, la tâche des participants était de juger si l'intonation de la réponse était congruente ou non par rapport à la question posée. Ils répondaient « oui » ou « non », en appuyant sur l'un des deux boutons de réponse. Pour un participant donné, l'ordre de présentation des dialogues était pseudo-aléatoire (pas plus de 3 dialogues appartenant à la même condition expérimentale présentés de façon successive), et l'ordre de présentation était différent d'un participant à un autre.

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 [11]. La référence commune est placée sur le mastoïde gauche. 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 150 ms avant la présentation du mot cible et continue pendant 2200 ms, avec une fréquence d'échantillonnage de 250 Hz. Les essais sont ensuite moyennés afin d'obtenir les PEs, après que ceux comportant des mouvements oculaires aient été rejetés. En outre, pour chaque phrase réponse, deux types d'enregistrement ont été réalisés : un à partir du mot médian et un autre à partir du mot terminal. Il est ainsi possible d'analyser les PEs, non seulement lorsque le mot cible porte un accent focal (congruent vs incongru), mais également lorsqu'il n'en porte pas (non attendu vs attendu).

Les différences entre conditions expérimentales ont ensuite été testées statistiquement au moyen d'analyses de variance (i.e., ANOVAs) effectuées sur l'amplitude moyenne des données enregistrées au niveau des électrodes médianes, dans des rang de latence déterminés après inspection visuelle des PEs. Pour toutes les ANOVAs effectuées, les facteurs suivants ont été utilisés: Congruence du focus (M+ vs M-, T+ vs T-), Présence du focus (M+ vs M0, T+ vs T0), électrodes (4).

3. RÉSULTATS

3.1. Données comportementales

Par rapport à la question présentée au début de chaque dialogue, les participants ont jugés 95% des réponses cohérentes comme étant acceptables, et 92% des réponses incohérentes comme n'étant pas acceptables.

Ainsi Ces données démontrent clairement que les participants étaient sensibles aux accents focaux incongrus.

3.2. Données électrophysiologiques

Les principaux résultats sont résumés sur la figure 1.

Mot Médian

Les résultats de l'analyse de variance, effectuée entre 300 et 1000 ms, ont révélé que lorsque les mots Médians sont prosodiquement incongrus, soit parce qu'ils portent un accent focal non attendu (condition M-T0), soit parce qu'ils ne portent pas d'accent focal alors que l'on en attendait un (condition M0T-), ils suscitent l'occurrence d'une composante positive plus ample [$F(1,15) = 15.76$; $p = 0.001$] que lorsque les mots Médians sont prosodiquement congruents, soit parce qu'ils portent un accent focal attendu (condition M+T0), soit parce qu'ils ne portent pas d'accent focal, mais qu'il n'était pas attendu (condition M0T+). En revanche, les mots porteurs d'un accent focal bien placé et les mots non focalisés pour lesquels aucun n'accent focal n'était attendu ne sont

pas significativement différents.

Mot Terminal

Les résultats de l'analyse de variance, effectuée entre 200 et 1000 ms, ont révélés que lorsque les mots Terminaux sont prosodiquement incongrus (conditions T0M- et T-M0), ils suscitent l'occurrence d'une composante négative plus ample [$F(1,15) = 29.15$; $p < 0.001$] que lorsque les mots terminaux sont prosodiquement congruents (conditions T+M0 et T0M+). De plus, comme pour les mots médians, les mots porteurs d'un accent focal bien placé et les mots non focalisés pour lesquels aucun n'accent focal n'était attendu ne sont pas significativement différents.

4. DISCUSSION

Les résultats de notre expérience démontrent clairement que dans une situation de dialogue, les patrons prosodiques sont traités en temps réel. De plus, des différences claires ont été trouvées entre les patrons prosodiques qui sont cohérents avec les contraintes pragmatiques imposées par la structure informationnelle du dialogue (conditions M+T0 et

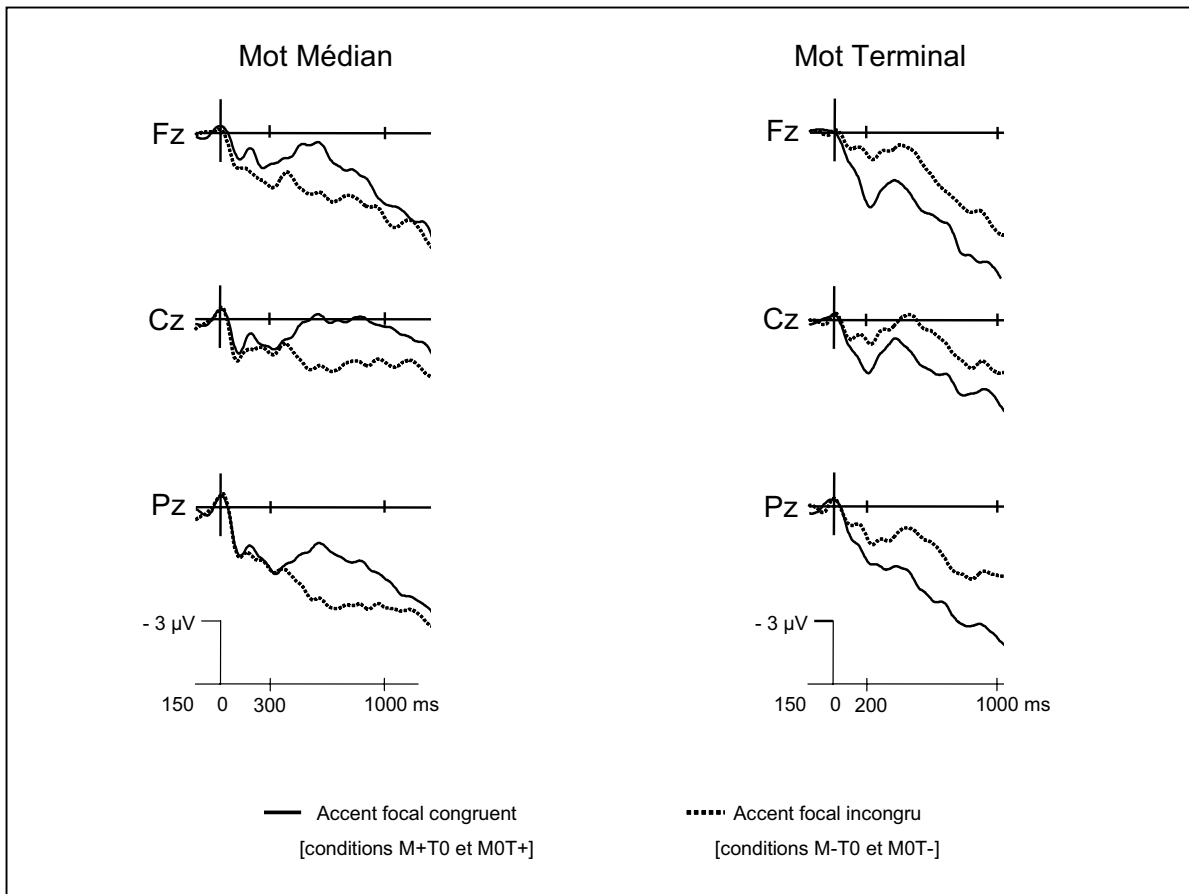


Figure 1: Grand-moyennage des données enregistrées chez 16 participants. Potentiels évoqués associés à la présentation du mot médian (colonne de gauche) ou terminal (colonne de droite) de la réponse, lorsqu'il porte un accent focal congruent (tracé plein) ou incongru (tracé en pointillé). Les données présentées sont enregistrées au niveau des électrodes frontale (Fz), 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.

M0T+), et les mots pragmatiquement incongrus (conditions M-T0 et M0T-). Cet effet d'incongruité pragmatique est largement modulé par la position de l'accent focal dans la réponse. Ainsi, lorsque les incongruités pragmatiques sont situées au milieu de la phrase, elles suscitent l'occurrence d'une composante positive de type P300. Cette composante est connue dans la littérature comme reflétant le traitement d'un événement surprenant [12]. Ainsi, un mot pragmatiquement incongru dans la réponse, soit parce qu'il porte un accent focal incongru, soit parce qu'il ne porte pas d'accent focal alors qu'on en attend un, provoque un effet de surprise chez l'auditeur, qui se reflète dans son activité électrique cérébrale par l'occurrence d'une composante P300.

En revanche, lorsque les incongruités pragmatiques sont situées en fin de phrase, elles suscitent l'occurrence d'une composante négative de type N400. Cette composante est connue dans la littérature comme reflétant le traitement sémantique de mots qui ne sont pas attendus dans le contexte d'un autre mot, d'une phrase, ou d'un contexte plus large tel qu'un dialogue [13]. Plus directement relié à nos résultats, cette composante est également présente en fin de phrase, lorsque l'auditeur éprouve des difficultés à intégrer l'ensemble des informations disponibles (e.g., syntaxique, sémantique et pragmatique) pour se construire une représentation cohérente de l'énoncé [14]. Ainsi, contrairement à l'aspect surprenant d'un focus pragmatiquement incongru en milieu de phrase, qui peut éventuellement être récupéré en fin de phrase grâce aux informations fournies par la suite du message, un focus pragmatiquement incongru en fin de phrase ne peut pas être récupéré par la suite du message. Par conséquent, l'auditeur ne bénéficie d'aucune information supplémentaire lui permettant éventuellement de pallier le problème posé par l'accent focal mal placé, soit sur le mot médian (conditions M-T0), soit sur le mot terminal (conditions M0T-). Cette difficulté d'intégration serait reflétée dans les potentiels évoqués par l'occurrence d'une composante N400.

En conclusion, ces résultats sont importants pour au moins deux raisons. Ils démontrent, en effet, que le traitement du focus prosodique contrastif est réalisé en temps réel par l'auditeur afin de construire une représentation cohérente du dialogue. Ils complètent également ceux précédemment obtenus en psycholinguistique et acoustique, montrant l'importance de l'accentuation et de l'intonation dans le traitement du langage parlé au niveau du discours. Il sera par la suite intéressant d'étudier dans quelle mesure les effets observés sont dépendants de la tâche utilisée (i.e., explicite sur l'intonation), et déterminer ainsi s'il sont reproductibles lorsque l'attention de l'auditeur est focalisée sur un autre aspect que l'intonation, comme dans une tâche de jugement sémantique (i.e., implicite sur l'intonation)

BIBLIOGRAPHIE

- [1] P. Warren. Prosody and Parsing: An Introduction. *Lang. and Cog. Proc.*, 11:1-16, 1996.
- [2] H. Berger. Uber das Electrenkephalogramm das menchen. *Archiv für Psychiatrie*, 87:527-570, 1929.
- [3] M. Kutas and S.A Hillyard. Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207:203-205, 1980.
- [4] L. Osterhout and P.J. Holcomb. Event-Related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, 31:785-804, 1992.
- [5] K.B.E. Böcker, M.C.M. Bastiaansen, J. Vroomen, C.H.M. Brunia and B. De Gelder. An ERP correlate of metrical stress in spoken word recognition, *Psychophysiology*, 36:706-720, 1999.
- [6] K. Steinhauer, K. Alter and A.D. Friederici. Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience*, 2 :191-196, 1999.
- [7] A. Schirmer , S.A. Kotz and A.D. Friederici. Sex differentiates the role of emotional prosody during word processing. *Cognitive Brain Research*, 14:228-233, 2002.
- [8] C. Astésano, M. Besson and K. Alter. Brain potentials during semantic and prosodic processing in French. *Cognitive Brain Research*, 18:172-184, 2003.
- [9] K. Lambrecht. *Information Structure and Sentence Form*. Cambridge, Cambridge University Press. 1994.
- [10] C. Astésano, C. Magne, M. Morel, A. Coquillon, R. Espesser, M. Besson and A. Lacheret. Marquage acoustique du focus contrastif non codé syntaxiquement en français.
- [11] H. Jasper. The 10-20 electrode system of the International Federation. *Electroencephalographie and Clinical Neurophysiology*, 10:370-375, 1958.
- [12] E. Donchin. Surprise! . . . Surprise? *Psychophysiology*, 18:493-513, 1981.
- [13] J.J.A. van Berkum , P. Zwitserlood , P. Hagoort and C. M. Brown. When and how do listeners relate a sentence to the wider discourse? Evidence from the N400 effect. *Cognitive Brain Research*, 17:701-718, 2003.
- [14] P. Hagoort. Interplay between Syntax and Semantics during Sentence Comprehension: ERP Effects of Combining Syntactic and Semantic Violations. *J. of Cog. Neurosc*, 15:883-899, 2003.