



L'information accentuelle est-elle représentée dans le lexique mental des locuteurs du français ?

Amandine Michelas, Sophie Dufour
Aix Marseille Univ, CNRS, LPL, Aix-en-Provence, France

michelas@lpl-aix.fr, sophie.dufour@lpl-aix.fr

RESUME

Une particularité prosodique du français est que l'accent primaire ne permet pas de distinguer deux mots de sens différents. À l'aide d'un paradigme d'amorçage de répétition, nous avons examiné si l'information accentuelle est représentée dans le lexique mental des français. En comparaison à une condition contrôle, nous avons observé une diminution des temps de réaction, que le mot cible « ban'**deau** » accentué sur sa syllabe finale soit précédé de l'amorce « ban'**deau** » elle-même accentuée sur sa syllabe finale ou qu'il soit précédé de l'amorce « bandeau » inaccentuée. Un tel résultat suggère que les mots sont stockés indépendamment de la présence ou non d'un accent en français. D'un point de vue théorique plus général, ce résultat est compatible avec les modèles abstractionnistes de la reconnaissance des mots parlés dans lesquels les informations acoustiques non pertinentes à l'identification des mots seraient écartées du signal de parole avant l'accès au lexique.

ABSTRACT

Is stress information represented in the mental lexicon of French speakers?

One striking characteristic of French is that stress does not allow distinguishing two words of different meanings. Using the repetition priming paradigm, we examined whether stress information is represented in the mental lexicon of French speakers. In comparison to a control condition, we observed shorter reaction times both when the target word “ban'**deau**” stressed on its final syllable was preceded by the prime “ban'**deau**” also stressed on its final syllable and when it was preceded by the word “bandeau” unstressed. Such a result suggests that words are stored independently of the presence or absence of stress in French. At a more theoretical level, this result is compatible with abstract models of spoken word recognition in which acoustic details irrelevant to identification are discarded from the speech signal before lexical access.

MOTS-CLES : Perception de la parole, reconnaissance des mots, représentations abstraites, prosodie du français, accent primaire

KEYWORDS: Speech perception, word recognition, abstract representations, French prosody, stress.

1 Introduction

Une des caractéristiques prosodiques les plus frappantes du français par rapport aux autres langues romanes est le fait que l'accent primaire ne permet pas de distinguer deux mots de sens différents comme c'est le cas en espagnol (ex. **'bebe** « elle boit » vs. be**'be** « bébé »). Au contraire, en français l'accent primaire¹ affecte une unité plus grande que le mot, le syntagme accentuel, et c'est toujours la dernière syllabe du syntagme accentuel qui porte l'accent. Cet accent est obligatoire et est caractérisé par deux corrélats acoustiques principaux : l'allongement de la durée de la dernière syllabe du syntagme et une montée de fréquence fondamentale (f0) associée à cette syllabe lorsque le syntagme n'est pas en position finale d'énoncé (Jun & Fougeron, 2000 ; Welby, 2006). Le fait qu'en français l'accent affecte toujours la dernière syllabe du syntagme accentuel a une conséquence importante : un même mot peut-être accentué ou non en fonction de sa position au sein du syntagme. Par exemple, le mot BANDEAU est accentué dans la phrase « On m'a parlé [d'un petit BAN'**DEAU**]_{SA} » car il est situé en position finale du syntagme accentuel (SA). Au contraire, ce même mot n'est pas accentué dans la phrase « On m'a parlé [d'un bandeau '**rouge**]_{SA} » car il n'est pas en position finale du syntagme accentuel. De ce fait, même si l'accent ne permet pas de créer de contrastes de sens au niveau du mot, les français sont confrontés quotidiennement aux versions accentuées et inaccentuées des mots. Dans cette étude, nous nous sommes donc demandé si l'accent est intégré ou non aux représentations mentales des mots chez les français comme cela a été montré chez les auditeurs espagnols (Soto-Faraco et al., 2001).

Au regard des modèles théoriques de la reconnaissance des mots parlés, seuls les modèles exemplaristes dans lesquels chaque mot est associé à de multiples *tokens* encodant des détails acoustiques fins (Goldinger, 1998) postulent que l'information accentuelle est stockée dans le lexique mental des français. Au contraire, les modèles abstractionnistes (McClelland & Elman, 1986), supposant une première étape de normalisation du signal de parole au cours de laquelle les détails acoustiques fins non pertinents pour l'identification des mots seraient écartés du signal de parole, postulent que la présence ou non d'un accent primaire chez les locuteurs du français ne serait pas intégrée à leurs représentations lexicales.

Dans la présente étude, nous avons cherché à départager ces deux types de modèles en utilisant un paradigme d'amorçage de répétition à long-terme. Ce paradigme consiste à présenter dans un premier temps un bloc de mots (bloc amorce) aux participants sur lesquels ils doivent réaliser une tâche (i.e., décision lexicale). Dans un second temps, un second bloc de mots (bloc cible) leur est présenté, la moitié des mots ayant déjà été rencontrée dans le premier bloc, l'autre n'ayant jamais été rencontrée. Typiquement, les mots qui sont répétés sont reconnus plus rapidement que les mots non répétés. L'atténuation de cet effet d'amorçage de répétition lors de la modification d'une dimension particulière entre les deux blocs (par exemple lors de la modification de la présence/absence d'un accent), indiquerait en accord avec les modèles à exemplaires que le même mot prononcé avec une différence d'accent activerait différentes représentations lexicales et que des spécificités liées à la présence ou non d'un accent primaire seraient stockées en mémoire. Au contraire, aucune modulation de l'effet d'amorçage en fonction de la présence/absence d'un accent

¹ Notons qu'il existe également un accent secondaire en français, qui est optionnel, qui affecte le début du syntagme accentuel et qui n'a pas les mêmes corrélats acoustiques que l'accent primaire. Contrairement à l'accent primaire, l'accent secondaire n'est pas accompagné d'un allongement de la syllabe. Il est seulement associé à une montée de f0 qui affecte le début du premier mot de contenu du syntagme accentuel (Welby, 2006).

indiquerait en accord avec les théories abstractionnistes de la reconnaissance des mots parlés que le même mot prononcé avec des différences accentuelles activerait la même représentation lexicale. Des participants de langue maternelle française ont eu à réaliser une tâche de décision lexicale, dans laquelle ils devaient décider le plus rapidement et le plus précisément possible si les stimuli présentés constituaient ou non un mot de la langue française. De façon à nous assurer de l'origine lexicale de l'effet d'amorçage de répétition susceptible d'être observé sur les mots, les pseudo-mots pouvaient être répétés soit avec une accentuation identique ou soit avec une accentuation différente. La logique sous-jacente à une telle manipulation est que des effets d'origine lexicale devraient être plus importants pour les mots que pour les pseudo-mots, ces derniers par définition n'étant pas associés à une représentation lexicale.

2 Méthode

2.1 Participants

48 participants de langue maternelle française ont pris part à l'expérience. Tous les participants étaient des étudiants d'Aix-Marseille Université et ont tous rapporté n'avoir aucun trouble de l'audition, de la parole ou trouble neurologique.

2.2 Matériel

48 mots cibles bisyllabiques de structure CVCV ont été sélectionnés sur la base de données Lexique.org (New et al., 2005) et ont été utilisés à la fois comme amorces et comme cibles. La fréquence moyenne de ces mots était de 7,60 par million d'occurrences. De plus, 48 pseudo-mots cibles bisyllabiques de structure CVCV, servant également comme amorces, ont été créés en changeant le dernier phonème de mots réels (ex. « baisi » créé à partir du mot « baiser »). 16 autres mots utilisés comme amorces contrôles ont été également sélectionnés et 16 autres pseudo-mots utilisés également comme amorces contrôles ont été créés.

Afin d'obtenir les versions accentuées et inaccentuées de chacun des stimuli, nous avons demandé à une locutrice de langue maternelle française de prononcer les 64 mots et 64 pseudo-mots au sein de phrases porteuses dans lesquelles ils pouvaient être accentués ou non en fonction de leur position dans la phrase (ex. On m'avait parlé [d'un bandeau '**bleu**]_{SA} qui était joli vs. On m'avait parlé [d'un petit ban'**deau**]_{SA} qui était joli ; ex. On m'avait parlé [d'un baisi '**bleu**]_{SA} qui était joli vs. On m'avait parlé [d'un petit bai'**si**]_{SA} qui était joli). Afin d'éviter les effets de coarticulation lié à la parole en contexte, chaque mot a d'abord été extrait de sa phrase porteuse, puis présenté auditivement à la locutrice. Celle-ci devait répéter chacun des mots et des pseudo-mots dans leurs versions accentuées et inaccentuées. Les 128 répétitions ainsi obtenues ont été enregistrées à une fréquence d'échantillonnage de 44 100 Hz, segmentées puis normalisées en intensité à un niveau de 70 dB SPL.

Des analyses acoustiques ont été ensuite conduites grâce au logiciel Praat (Boersma & Weenink, 2015) sur ces 128 répétitions afin de s'assurer que celles-ci étaient produites avec les patrons accentuels attendus. La figure 1 illustre le mot « bandeau » et le pseudo-mot « baisi » produits dans leurs versions accentuées et inaccentuées.

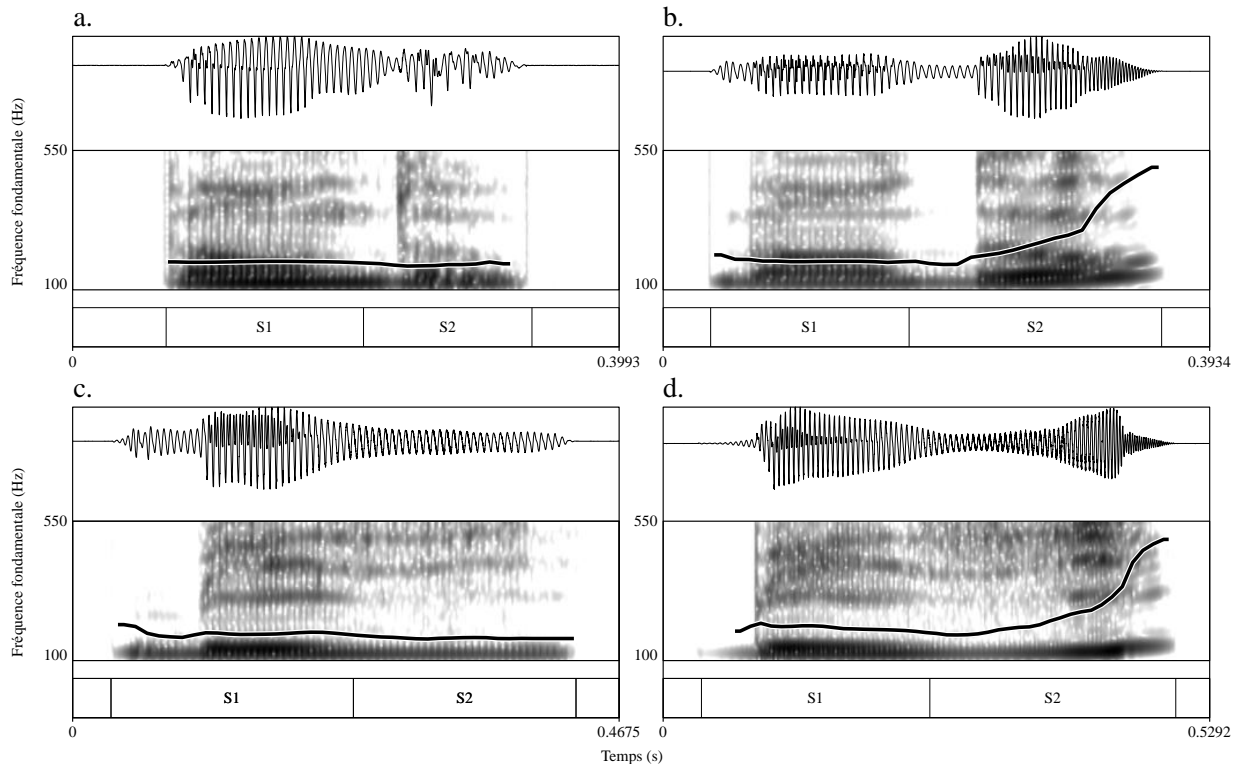


FIGURE 1. Onde sonore, spectrogramme et courbe de f0 pour le mot « bandeau » dans sa version inaccentuée (a) et accentuée (b) et le pseudo-mot « baisi » dans sa version inaccentuée (c) et accentuée (d).

L'allongement de la syllabe finale de chaque stimulus ainsi que la montée de f0 associée à cette syllabe finale ont été mesurés (cf. Table 1).

		Version inaccentuée		Version accentuée	
		1 ^{ère} syllable du mot	2 ^{ème} syllable du mot	1 ^{ère} syllable du mot	2 ^{ème} syllable du mot
Mots	Durée syllabique (ms)	144	143	142	239
	Minimum de f0* (Hz)	187	176	182	171
	Maximum de f0* (Hz)	181	174	178	402
	Montée de f0 (%)	-3	-1	-2	136
Pseudo-mots	Durée syllabique (ms)	153	153	153	242
	Minimum de f0* (Hz)	185	180	184	180
	Maximum de f0* (Hz)	183	178	183	374
	Montée de f0 (%)	-1	-1	-1	107

TABLE 1. Caractéristiques acoustiques des mots et pseudo-mots cibles dans leurs versions accentuées et inaccentués. * Pour les versions inaccentuées des mots et pseudo-mots, les valeurs minimum et maximum de f0 correspondent aux valeurs de début et de fin du plateau de f0 observé.

Comme attendu, les 48 mots cibles ainsi que les 48 pseudo-mots cibles avaient une syllabe finale plus longue ($t(47)=17.73$, $p<.0001$) et plus haute que la première syllabe ($t(47)=21.26$, $p<.0001$) uniquement dans leurs versions accentuées. Les mêmes vérifications ont été effectuées pour les 16 mots et les 16 pseudo-mots utilisés comme amorce contrôle.

Deux blocs de stimuli ont déjà été créés. L'un était utilisé comme bloc amorce et l'autre comme bloc cible. A l'intérieur de chaque bloc, la moitié des stimuli était accentuée et l'autre moitié était inaccentuée. Le bloc cible était constitué des 48 mots cibles et des 48 pseudo-mots cibles. Parmi les mots et les pseudo-mots cibles, 16 ont été utilisés dans la condition « répétée accentuation identique », 16 dans la condition « répétée accentuation différente » et 16 dans la condition contrôle (i.e. non répétée). Le bloc amorce était également constitué de 48 mots et de 48 pseudo-mots. Parmi les mots et les pseudo-mots, 16 consistaient en la répétition des cibles avec accentuation identique, 16 en la répétition des cibles avec une accentuation différente et les 16 autres correspondaient aux amorces contrôles.

Afin que chaque cible (mots, pseudo-mots) soit entendue dans chaque condition d'amorçage (répétée accentuation identique, répétée accentuation différente, contrôle) et que chaque participant n'entende qu'une seule fois le même mot cible, trois listes expérimentales ont été créées. Les trois listes ont ensuite été divisées en deux sous-listes de manière à ce que chaque stimulus soit entendu dans sa version accentuée et non accentuée.

2.3 Procédure

Les participants, munis d'un casque audio, ont été testés individuellement dans une chambre insonorisée et les stimuli leur étaient présentés à un niveau sonore confortable. La présentation des stimuli était contrôlée par un ordinateur grâce au logiciel E-Prime (version 2.0, Psychology Software Tools). Pour chaque stimulus du bloc amorce et du bloc cible, les participants devaient indiquer le plus rapidement et le plus précisément possible s'il s'agissait d'un mot de la langue française ou non en fournissant la réponse « mot » avec leur main dominante. Les temps de réponse (TRs) étaient enregistrés à partir du début des stimuli. A l'intérieur de chaque bloc, les stimuli étaient présentés dans un ordre aléatoire. La réponse du participant et le début de présentation du stimulus suivant étaient séparés par un délai de 2 secondes. Les participants ont été testés sur une seule des sous-listes expérimentales et ont commencé l'expérience avec 12 essais d'entraînement.

3 Résultats

Uniquement les temps de réaction obtenus dans le bloc cible ont été analysés. Six items engendrant des taux d'erreurs supérieurs à 40% ont été exclus des analyses. Une donnée aberrante correspondant à un TR de 372ms ainsi que 25 données aberrantes correspondants à des TRs supérieurs à 2000ms ont été supprimées. Les analyses ont donc portées sur 3973 données. Les temps de réaction (TRs) moyens ainsi que les pourcentages de réponses correctes sont présentés dans la Figure 2.

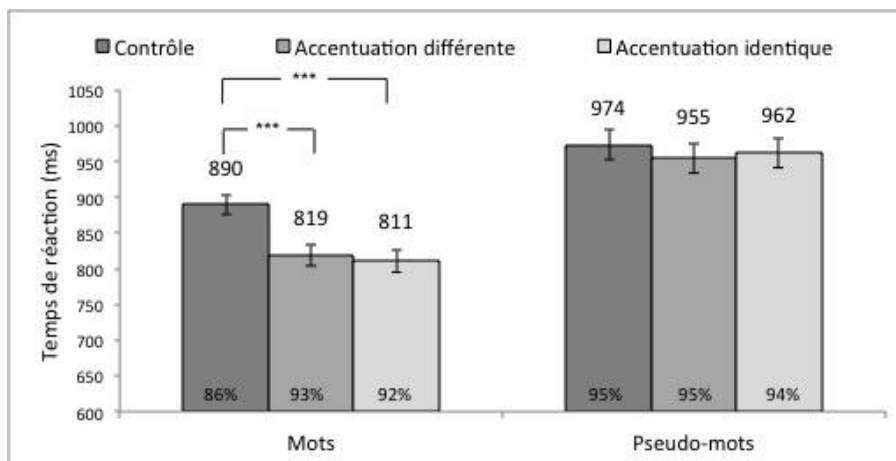


FIGURE 2. Temps de réaction moyen (en ms) et pourcentage de réponses correctes obtenus dans le bloc cible en fonction de la lexicalité et du type d'amorce (les barres représentent les erreurs standards).

Afin d'analyser ces temps de réaction, nous avons utilisé un modèle linéaire à effets mixtes sur les transformations logarithmiques des TRs (lme4 package in R-studio statistics Version 0.99.903). Ce modèle avait pour effets fixes le type d'amorce (contrôle/répétée accentuation identique/répétée accentuation différente) et la lexicalité (mot/pseudo-mot). La structure aléatoire du modèle incluait des pentes différentes pour les participants pour le facteur lexicalité et pour le facteur type d'amorce, des pentes différentes pour les mots pour le facteur type d'amorce ainsi que des interceptes différents pour les participants et les mots (Barr et al., 2013). Le modèle a révélé un effet principal significatif du type d'amorce ($F=34.27$, $p<.0001$) ainsi que de la lexicalité ($F=59.02$, $p<.0001$). L'interaction type d'amorce * lexicalité était également significative ($F=17.52$, $p<.0001$). Afin de décomposer cette interaction, le design 2x3 a été converti en un design 1x6. Les comparaisons multiples ont été obtenues grâce à la fonction glht du package multcomp (Bretz et al., 2011) avec des p-values ajustées avec une correction de Bonferroni.

Les temps de réponse étaient plus rapides pour les mots cibles précédés d'une amorces répétée avec accentuation identique (811ms) que pour les mots cibles précédés d'une amorce répétée avec accentuation différente (819ms ; $z=-8.82$; $p<.0001$). Ils étaient également plus rapides pour les mots cibles précédés d'une amorce répétée avec accentuation identique (811ms) que pour les mots cibles précédés d'une amorce répétée avec accentuation différente (819ms ; $z=1.13$; $p>.20$). Nous avons également observé aucune différence entre les pseudo-mots quelque soit le type d'amorce qui était présenté (pseudo-mots accentuation identique vs. pseudo-mots contrôles : $z=-1.15$, $p>.20$; pseudo-mots accentuation différente vs. pseudo-mots contrôles : $z=-2.16$, $p>.20$; pseudo-mots accentuation identique vs. pseudo-mots accentuation différente: $z=-1.00$, $p>.20$).

Les taux de réponses correctes ont été analysés à l'aide d'un modèle logit à effets mixtes. Ce modèle avait pour effets fixes le type d'amorce (contrôle/répétée accentuation identique/répétée accentuation différente) et la lexicalité (mot/pseudo-mots). La structure aléatoire du modèle incluait un intercepte différent pour les participants et les mots mais pas de pentes différentes car l'inclusion de celles-ci ne permettait pas au modèle de converger (Barr et al., 2013). Les effets globaux ont été obtenus grâce à la fonction afex::mixed (Sinnmann et al., 2013). Le modèle a révélé un effet principal significatif de la lexicalité ($X^2=7.35$, $p<.01$). L'effet du type d'amorce échouait à atteindre

la significativité ($X^2=5.60$, $p=.06$), alors que l'interaction type d'amorce * lexicalité était significatif ($X^2=9.43$, $p<.01$). Afin de décomposer cette interaction, le design 2x3 a été converti en un design 1x6. Les comparaisons multiples ont été obtenues grâce à la fonction `glht` du package `multcomp` (Bretz et al., 2011) avec des p-values ajustées avec une correction de Bonferroni. Le modèle a révélé plus de réponses correctes pour les mots cibles précédés d'une amorce répétée avec accentuation identique (92%) que pour les mots cibles précédés d'une amorce contrôle (86% ; $z=-3.31$; $p<.05$). Les participants ont également donné plus de réponses correctes lorsque les mots cibles étaient précédés d'une amorce répétée avec accentuation différente (93%) que lorsque les mots cibles étaient précédés d'une amorce contrôle (86% ; $z=-3.83$, $p<.01$). Au contraire, aucune différence de performance n'a été observée entre les mots cibles précédés d'une amorce répétée avec accentuation identique (92%) et les mots cibles précédés d'une amorce répétée avec accentuation différente (93% ; $z=-0.56$; $p>.20$). Aucune différence significative n'a été observée pour les pseudo-mots quelque soit le type d'amorce qui était présenté ($p>.20$).

4 Discussion

L'hypothèse sous-jacente à notre recherche était que si les indices liés à la présence/absence d'un accent sont intégrés au lexique mental des locuteurs du français, une atténuation de l'effet d'amorçage de répétition devrait être observée lorsque des amorces et des cibles diffèrent du point de vue de leur accentuation. Au contraire, si les indices liés à la présence/absence d'un accent ne sont pas intégrés aux représentations lexicales des locuteurs du français, aucune modulation dans l'effet d'amorçage de répétition ne devrait être observée lorsque des amorces et des cibles diffèrent sur leur accentuation. Les résultats obtenus dans cette étude ne montrent aucune diminution de la taille de l'effet d'amorçage de répétition lors d'un changement d'accentuation entre les amorces et les cibles. En effet, les temps de réponse des participants étaient plus rapides pour les mots cibles précédés d'une amorce répétée avec accentuation identique que pour les mots cibles précédés d'une amorce contrôle. Les temps de réponse étaient également plus rapides pour les mots cibles précédés d'une amorce répétée avec accentuation différente que pour les mots cibles précédés d'une amorce contrôle. Cependant, de façon cruciale, aucune différence de temps de réaction n'a été observée entre les mots cibles précédés d'une amorce répétée avec accentuation identique et les mots cibles précédés d'une amorce répétée avec accentuation différente. Une telle observation plaide en faveur de l'existence de représentations n'encodant pas l'information accentuelle chez les locuteurs du français et indique qu'un même mot prononcé avec ou sans accent active la même représentation lexicale de base.

L'avantage majeur du paradigme d'amorçage de répétition à long-terme que nous avons utilisé est qu'il nous permet de sonder la nature des représentations lexicales et plus particulièrement quel type d'information elles sont susceptibles d'encoder. L'origine lexicale de nos effets est attestée par l'absence d'effet d'amorçage de répétition avec des pseudo-mots. Les pseudo-mots, n'étant par définition associés à aucune représentation lexicale, ne sont pas réactivés à un niveau lexical de traitement lors de leur seconde présentation et aucun effet d'amorçage de répétition n'a ainsi été observé. Dans deux études précédentes (Michelas et al. 2016 ; Michelas et al., 2017) et à l'aide d'une tâche « même-différent », nous avons montré que les auditeurs du français étaient capables de traiter la différence entre des mots accentués et des mots inaccentués et ceci à un niveau plus abstrait de traitement que de simples distinctions acoustiques contrairement à ce qui avait été suggéré dans des précédentes études (Dupoux et al., 1997). Toutefois même si les français sont capables de percevoir les différences entre un mot inaccentué et accentué, la présente étude montre que l'information accentuelle n'est en aucun cas encodée dans leurs représentations lexicales.

D'un point de vue théorique plus général, un tel résultat s'intègre dans le débat actuel concernant le format des représentations lexicales et en particulier sur l'inclusion ou non de variations accentuelles dans ces représentations. Comme nous l'avons vu précédemment, seuls les modèles exemplaristes de la reconnaissance des mots parlés postulent que l'information accentuelle est intégrée aux représentations lexicales des locuteurs du français. Néanmoins, une façon de rendre compte de nos résultats dans un cadre exemplariste est d'envisager que toutes les variations acoustiques ne sont pas encodées dans les représentations lexicales et qu'un poids plus important serait accordé à certaines variations. Toutefois, le fait que nos manipulations de durée et de f0 n'aient pas eu d'impact sur la reconnaissance des mots semble être difficilement réconciliable avec une vision purement exemplariste de la reconnaissance des mots parlés. Nos résultats semblent donc être d'avantage en accord avec les modèles abstractionnistes tels que TRACE (McClelland & Elman, 1986) dans lesquels les détails acoustiques non pertinents pour identifier les mots en français sont écartés lors d'une première phase de normalisation du signal de parole et dans lesquels des représentations symboliques abstraites n'encodant pas de variation acoustique sont contactées.

Pour conclure, il apparaît donc que l'information accentuelle soit traitée et stockée différemment selon la structure prosodique des langues. Pour une langue telle que le français où l'information accentuelle n'est pas pertinente pour identifier deux mots de sens différents, les détails acoustiques liés à la présence ou non d'un accent seraient utilisés à un niveau pré-lexical de traitement (ex. pour segmenter mots ; Christophe et al., 2004) mais ne seraient pas intégrés aux représentations lexicales. Au contraire, pour des langues à accent lexical tel que l'anglais ou l'espagnol, les représentations mentales des mots incluraient des paramètres accentuels tels que la position de l'accent (Soto-Farraco et al., 2001).

Remerciements

Cette recherche, menée dans le cadre du Labex BLRI (ANR-11-LABX-0036) et l'Institut Convergence ILCB (ANR-16-CONV-0002), a reçu le soutien du gouvernement français, par le biais de l'Agence nationale de la Recherche (ANR) et l'Initiative d'Excellence d'Aix-Marseille Université (A*MIDEX).

Références

- BARR, D. J., LEVY, R., SCHEEPERS, C., TILY, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language* 68(3), 255-278.
- BOERSMA, P., WEENINK, D. (2015). *Praat. Doing phonetics by computer* (Version 5.4.01, 2015). Computer program: www.praat.org (Last viewed April 30, 2015).
- BRETZ, F., HOTHORN, T., WESTFALL, P. H. (2011). *Multiple comparisons using R*. Boca Raton: CRC Press.
- CHRISTOPHE, A., PEPPERKAMP, S., PALLIER, C., BLOCK, E., MEHLER, J. (2004). Phonological phrase boundaries constrain lexical access I. Adult data. *Journal of memory and language* 51(4), 523-547.

- DUPOUX, E., PALLIER, C., SEBASTIAN, N., MEHLER, J. (1997). A distressing “deafness” in French?. *Journal of Memory and Language* 36(3), 406-421.
- GOLDINGER, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological review*, 105(2), 251.
- JUN, S. A., FOUGERON, C. (2000). A phonological model of French intonation. In A. Botinis (Eds), *Intonation* (pp. 209-242). Springer, Dordrecht.
- MCCLELLAND, J. L., ELMAN, J. L. (1986). The TRACE model of speech perception. *Cognitive psychology* 18(1), 1-86.
- MICHELAS, A., FRAUENFELDER, U. H., SCHÖN, D., DUFOUR, S. (2016). How deaf are French speakers to stress?. *The Journal of the Acoustical Society of America* 139(3), 1333-1342.
- MICHELAS, A., ESTEVE-GIBERT, N., DUFOUR, S. (2017). On French listeners’ ability to use stress during spoken word processing. *Journal of Cognitive Psychology*, 1-9.
- SINGMANN, H., BOLKER, B., WESTFALL, J., AUST, F. (2015). *afex: Analysis of factorial experiments*. R package version 0.13–145.
- SOTO-FARACO, S., SEBASTIAN-GALLES, N., & CUTLER, A. (2001). Segmental and suprasegmental mismatch in lexical access. *Journal of Memory and Language* 45, 412-432.
- WELBY, P. (2006). French intonational structure: Evidence from tonal alignment. *Journal of Phonetics* 34(3), 343-371.