

Perception des proéminences accentuelles en français langue maternelle et langue étrangère : effet de l'intonation et de la composante lexicale

Lorraine Baqué¹, M^a Ángeles Barquero^{1,2}, Marta Estrada¹
& Sandra Schwab²

¹ Departament de Filologia Francesa i Romànica
Universitat Autònoma de Barcelona

² École de langue et de civilisation françaises
Université de Genève

<lorraine.baque@uab.cat, angelesbarquero@hotmail.com,
marta.estrada@uab.cat, Sandra.Schwab@unige.ch>

Abstract

This research aims at examining the perception of accentual groups in French by French-speaking listeners and by Spanish/Catalan-speaking learners of French. The material was composed of French five-syllable sequences produced by native speakers of French and by Spanish/Catalan-speaking learners of French. The sequences were produced in the assertive and interrogative modalities. Moreover, a delexicalized version of the sequences was created. Participants listened to each sequence and had to decide whether it was composed of 1 or 2 groups. Results show that the listeners were more likely to perceive 2 groups when 1) the sequences were produced by learners of French, 2) the sequences were presented in the delexicalized version, and 3) the sequences' speech rate decreased.

Mots-clés : perception, groupes accentuels, proéminences, acquisition d'une langue seconde, espagnol, français

1. Introduction

L'accent primaire du français et de langues comme l'espagnol et le catalan diffère par sa fonction (démarcative *vs* culminative et distinctive), par sa position (oxytone *vs* libre), par son domaine d'application (de groupe *vs* lexical) ainsi que par les marques acoustiques qui le manifestent (durée *vs* fréquence fondamentale + durée/intensité, cf. Llisterri, Machuca, de la Mota, Riera & Ríos, 2005).

En français, la réalisation effective d'un accent sur les mots accentuables dépend, entre autres, du débit (Martin 2006, Delais-Roussarie & Post 2008), de la modalité (Wunderly 1988) et du nombre de syllabes du groupe (Rossi 1999). Ainsi, par exemple, la probabilité d'apparition d'accents augmenterait avec la réduction du débit, en

modalité assertive (*vs* interrogative) ou, par exemple, dans les syntagmes nominaux de 5 syllabes, pour des structures rythmiques de type 2+3 syllabes (ex. « les monts_{AC} Kitami », *vs* 3+2 : « les montagnes chinoises »).

Des études préalables portant sur les productions des étudiants hispanophones de français langue étrangère ont montré une tendance à accentuer la quasi-totalité des mots lexicaux comme en espagnol et à ne pas intégrer les règles de désaccentuation au sein du groupe en français (Barquero 2012, Schwab 2012). Néanmoins, il existe des indices que cette surabondance d'accents dans les productions des apprenants peut varier en fonction de l'intonation et, notamment, diminuer dans les interrogatives totales (Estrada 2003, Santiago & Delais-Roussarie 2012), alors que les accents seraient maintenus dans les productions des francophones (Di Cristo 1998).

Concernant la perception, la plupart des modèles (Archibald 1995, Drescher & Kaye 1990, Altmann & Vogel 2002, Peperkamp & Dupoux 2002) prédisent que les apprenants dont la langue maternelle se caractérise par un accent libre et distinctif (tels les hispanophones) présenteront une « surdité accentuelle » nulle ou inférieure à celle des francophones, dont la langue possède un accent fixe et non distinctif au niveau lexical (*a contrario* Astésano, Bertrand, Espesser, & Nguyen 2012).

Par ailleurs, il a été mis en évidence des différences significatives concernant la « surdité accentuelle » en fonction de si la tâche entraîne un traitement de bas niveau (acoustique) ou de plus haut niveau (lexical) (Dupoux, Peperkamp & Sebastián-Gallés 2001).

Néanmoins, la plupart des travaux sur la perception de l'accent en L2 se basent sur des tests portant sur des mots isolés produits par des natifs, et peu s'intéressent, à notre connaissance, à la perception dans des structures au-delà de l'unité lexicale ou à la perception de productions d'apprenants de L2.

Par conséquent, dans cette étude, nous nous sommes intéressées à la perception, par des francophones natifs et des apprenants hispano-catalanophones de français langue étrangère, des groupes accentuels dans des séquences produites par des francophones natifs et des apprenants de français langue étrangère. Nous avons également tenu compte de variables susceptibles d'avoir un impact sur la perception des groupes accentuels : la modalité, de la structure rythmique et du débit des productions ainsi que du type de parole.

2. Objectif

Cette étude vise à analyser dans quelle mesure les différences accentuelles entre le français et l'espagnol/catalan se traduisent

également par une perception différentielle des proéminences accentuelles en français. Le but est de vérifier si la perception dépend de la langue de l'auditeur, si la perception par des non-natifs de séquences produites par des pairs diffère de celle qu'ils ont des productions natives, et dans quelle mesure différents facteurs de variation, liés à la structure linguistique, aux propriétés acoustiques ou au type de traitement induit par la tâche affectent la perception des auditeurs natifs et non-natifs.

Ainsi, outre le groupe des auditeurs (Aud_Natif *vs* Aud_FLE) et des locuteurs (Loc_Natif *vs* Loc_FLE), trois facteurs sont pris en considération dans cette recherche : la modalité de la séquence (assertive *vs* interrogative totale), la structure rythmique de la séquence (2+3 *vs* 3+2 syllabes) et la présence ou absence d'information lexicale (séquences naturelles *vs* délexicalisées). Par ailleurs, l'effet du débit des séquences produites sera également analysé.

3. Méthode

3.1. Participants

Quatre groupes de sujets ont pris part à cette étude : deux groupes de locuteurs qui ont produit les séquences soumises ensuite au test de perception, et deux groupes d'auditeurs qui ont effectué ce test.

Les séquences ont été produites par 4 locuteurs francophones natifs (Loc_Natif, 2 hommes et 2 femmes d'âges compris entre 34 et 51 ans, moyenne = 41 ans) et 4 locuteurs hispano-catalanophones apprenants de français à Barcelone de niveau intermédiaire-avancé (Loc_FLE, 2 hommes et 2 femmes d'âges compris entre 21 et 22 ans).

Par ailleurs, 30 participants ont pris part au test de perception. Ils se divisent en deux groupes. Le premier groupe se compose de 15 francophones natifs (Aud_Natifs), d'âges compris entre 22 et 39 ans (moyenne de 31 ans) et le deuxième se compose de 15 hispano-catalanophones, apprenants de niveau intermédiaire (B1-B2 du CECRL) en français (Aud_FLE), d'âges compris entre 19 et 29 ans (moyenne de 22 ans).

3.2. Matériel

Le corpus utilisé dans cette étude se composait de quatre phrases issues d'un corpus plus vaste (DESACROM, Baqué & Estrada 2008).

- (1) Mon ami s'est disputé avec *la cousine du maire*
- (2) Ton mari a goûté *le fromage de chèvre*
- (3) Le concierge a volé *les clés du patron*
- (4) Le gouvernement a augmenté *le prix du tabac*

Les quatre phrases se terminaient par un syntagme nominal de 5 syllabes de type Dét+N1+du/de+N2. Dans deux des phrases, le premier substantif était bisyllabique (ex. cousine) et le deuxième monosyllabique (ex. maire), alors que dans les deux autres phrases c'était l'inverse : N1 monosyllabique (ex. clés) et N2 bisyllabique (ex. patron), ce qui correspond à des structures rythmiques potentielles de 2+3 syllabes *vs* 3+2 syllabes.

Les quatre phrases ont été produites en modalité assertive et en modalité interrogative totale par les quatre locuteurs francophones et les quatre locuteurs apprenants de français.

Le syntagme nominal final (ex. les clés du patron) a été extrait pour chacune des 64 productions (4 phrases x 2 modalités x 8 locuteurs). De plus, une version « délexicalisée » de ces séquences (i.e. syntagmes nominaux) a été créée en remplaçant toutes les consonnes par [s] et toutes les voyelles par [a] (à l'aide du script *Sasasa Delexicalizer* de Dellwo 2012).

Au total, 128 séquences en français ont donc été utilisées dans cette étude, 64 séquences dites « naturelles » et 64 dites « délexicalisées », la moitié produite par des natifs et l'autre moitié produite par des apprenants de français. Dans la moitié des séquences, les structures rythmiques potentielles correspondent à 2+3 syllabes et dans l'autre moitié à 3+2 syllabes.

3.3. Procédure

L'expérience de perception a été réalisée sur la plateforme Labguistic (Ménétreay & Schwab, sous presse ; www.labguistic.com) sur internet. Elle se composait de deux parties. Dans la première partie, les participants entendaient les séquences délexicalisées et dans la deuxième partie, ils entendaient les séquences naturelles. L'ordre de présentation des items était aléatoire et différent pour chaque participant. Après l'écoute de chaque séquence (5 écoutes maximum), les participants devaient indiquer si, selon eux, la séquence était formée d'un ou de deux groupes.

Avant de commencer l'expérience, les participants remplissaient un questionnaire biographique et linguistique. Au début de chaque partie, un bref entraînement permettait aux participants de se familiariser avec la tâche. L'expérience durait environ 20 minutes.

3.4. Analyse des données

Pour obtenir les valeurs de débit des différentes séquences produites par les locuteurs natifs et apprenants de FLE, nous avons aligné les différentes syllabes avec Praat (Boersma & Weeninck 2014), et divisé

la durée totale de la séquence par le nombre de syllabes prononcées pour obtenir un débit en syll/s.¹

Une première analyse des différences de débit en production a été menée à terme à l'aide d'un modèle de régression à effets mixtes (Baayen, Davidson & Bates 2008), avec pour variable dépendante le débit et pour variables aléatoires le Locuteur et l'Item. Les variables indépendantes introduites étaient le Groupe de locuteurs, la Modalité (assertive/interrogative), la Structure (2+3 syllabes/3+2 syllabes) et le Type de parole (naturelle/délexicalisée) ainsi que les possibles interactions.

Un deuxième modèle de régression logistique binaire à effets mixtes a été effectué pour analyser les variables qui incident sur la perception des items comme comportant un ou deux groupes (Variable « Réponse »). Les variables aléatoires étaient l'Auditeur et l'Item, avec les mêmes variables indépendantes que pour la première étude en plus du Groupe d'Auditeurs et du Débit, et des différentes interactions possibles.

Dans les deux cas, les variables et/ou interactions sans effet significatif sur la variable dépendante ont été supprimées du modèle.

4. Résultats et discussion

4.1. Analyse du débit produit

Une première analyse vise à rendre compte des différences de débit entre les items, ces différences étant susceptibles d'interférer avec les résultats de la perception des groupes accentuels (*cf.* section suivante).

Ainsi, comme on pouvait s'y attendre, comparées à celles des francophones, les productions des apprenants présentent globalement un débit plus lent (4.94 *vs* 6.19 syll/s., $F(1, 55) = 50.4998$, $p < .001$).

Malgré l'absence d'un effet simple de la modalité, il est intéressant de noter que les différences entre les francophones et les apprenants varient en fonction de la modalité ($F(1, 55) = 4.0112$, $p < .05$). Ainsi, comme l'illustre la Figure 1, comparées aux assertives, on constate dans les interrogatives une accélération du débit pour les Loc_Natif (6.34 *vs* 6.05 syll/s.) et un léger ralentissement chez les Loc_FLE : (4.89 *vs* 4.99 syll/s.).

¹ Aucune pause n'ayant été observée, le débit correspond également ici à la vitesse d'articulation en syllabes par seconde.

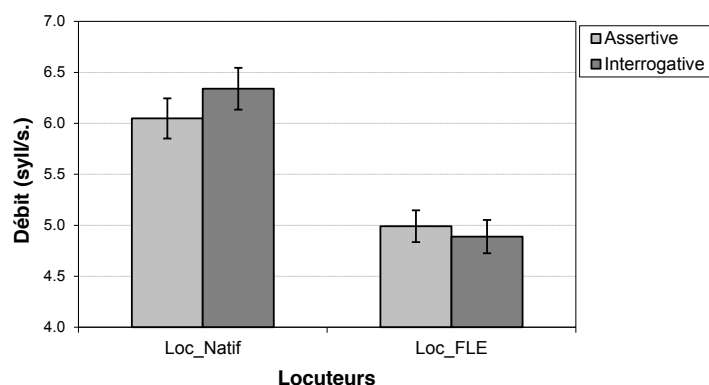


Figure 1 : Débit produit (syll/s.) en fonction du groupe de locuteurs et de la modalité. Les barres d'erreur représentent l'erreur standard de la moyenne

De plus, on remarque un effet simple de la structure sur le débit $F(1, 55) = 26.9793, p < .001$: la structure 2+3 présente un débit moyen supérieur à la structure 3+2 (6.04 vs 5.10 syll/s.). Notons également que cet effet n'est pas le même chez les francophones et chez les apprenants ($F(1, 55) = 14.4871, p < .001$) : les différences de débit dans les items produits par les Loc_Natif entre les structures 2+3 et 3+2 syllabes sont plus marquées (6.85 vs 5.54 syll/s.) que chez les Loc_FLE (5.22 vs 4.66 syll/s. respectivement), comme on peut le voir sur la Figure 2.

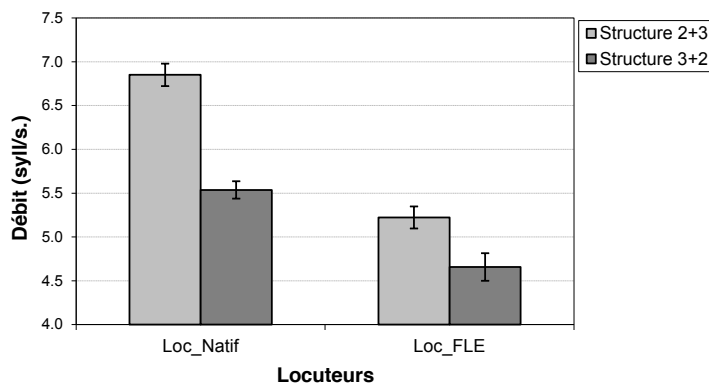


Figure 2 : Débit produit (syll/s.) en fonction du groupe de locuteurs et de la structure. Les barres d'erreur représentent l'erreur standard de la moyenne

4.2. Analyse de la perception des groupes accentuels

L'analyse du modèle de régression à effets mixtes que nous avons effectué montre tout d'abord un effet du groupe de locuteurs

(Loc_Natif *vs* Loc_FLE) sur le nombre de groupes perçus ($F(1, 3830) = 124.9558, p < .05$). Ainsi les items produits par des Loc_FLE sont plus souvent perçus comme comportant 2 groupes (46.87%) que ceux des Loc_Natif (40.07%).

De plus, le nombre de groupes perçus varie en fonction du type de parole ($F(1, 3830) = 6.5059, p < .001$) : la réponse « 2 groupes » est plus fréquente pour les items délexicalisés que pour les items de parole naturelle (46.63% *vs* 40.31%).

Enfin, il existe une relation entre le débit et le nombre de groupes perçus ($F(1, 3830) = 49.6974, p < .001$) : plus le débit est lent plus la probabilité de percevoir deux groupes augmente.

Par contre, nous n'observons pas d'effet de la modalité sur la probabilité de percevoir un ou deux groupes.

Par ailleurs, plusieurs interactions ont pu être mises en évidence. En premier lieu, bien qu'il n'y ait pas d'effet simple en fonction du groupe d'auditeurs ($F(1, 3830) = 2.5871, p < .100$), le nombre de groupes perçus varie en fonction du Groupe d'Auditeurs et du Type de parole ($F(1, 3830) = 10.8927, p < .001$). Ainsi, comme on peut le voir sur la Figure 3, les auditeurs francophones (Aud_Natif) perçoivent plus souvent 2 groupes dans les items délexicalisés (45.57%) que dans les items issus de parole naturelle (34.58%), tandis qu'il n'y a pas de différences entre les deux types de parole (47.68% *vs* 46.05%) pour les auditeurs apprenants (Aud_FLE).

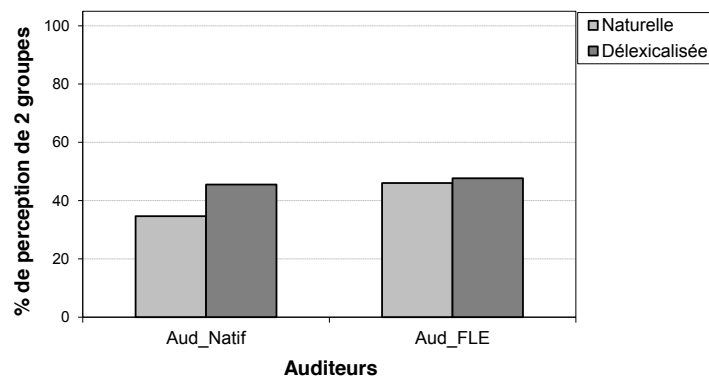


Figure 3 : Pourcentage de perception de 2 groupes en fonction du groupe d'auditeurs (Natif *vs* FLE) et du type de parole (Naturelle *vs* Délexicalisée)

Enfin, il existe également un effet d'interaction entre le groupe d'auditeurs et le débit ($F(1, 3830) = 13.7195, p < .001$), comme on peut le constater sur la Figure 4. Ainsi, si pour les deux types d'auditeurs, la probabilité de percevoir deux groupes augmente lorsque le débit

diminue, la relation est plus faible pour les Aud_FLE que pour les Aud_Natif.

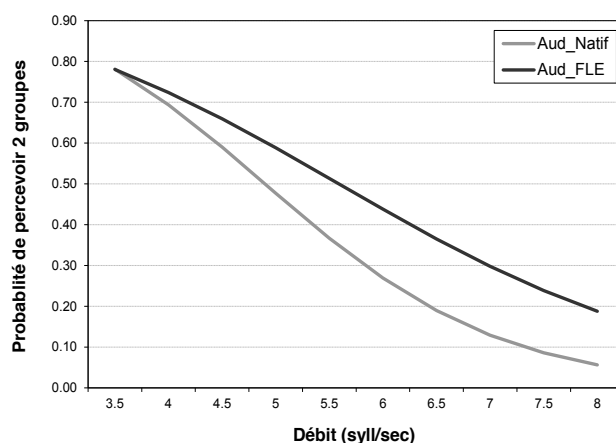


Figure 4 : Probabilité de percevoir 2 groupes en fonction du Groupe d'auditeurs (Natif vs FLE) et du débit de l'item (en syll/s.)

5. Discussion générale

L'analyse du débit des séquences produites en français par des locuteurs natifs et par des apprenants hispano-catalanophones de FLE de niveau intermédiaire-avancé confirme les études préalables qui font état d'une réduction du débit en langue étrangère. De plus, les apprenants semblent être moins sensibles aux autres facteurs (modalité et structure) que ne le sont les natifs, dont le débit s'accélère en modalité interrogative et pour la structure 2+3 syllabes. Les résultats concernant l'effet de la modalité pour les francophones correspondent aux attentes (*cf.* Introduction). Mais, pour ce qui est de la relation entre le débit et la structure, elle va dans le sens opposé des résultats de Rossi (1999). Une étude postérieure, axée sur la production, et contrôlant un éventuel effet de la structure syllabique, devrait nous permettre de mieux rendre compte de ce résultat.

Concernant le but principal de cette étude, c'est-à-dire l'analyse de la perception du nombre de groupes accentuels en français par des francophones natifs ou par des apprenants hispano-catalanophones de français langue étrangère, les résultats montrent en premier lieu que les productions des apprenants de FLE sont perçues comme étant plus souvent divisées en 2 groupes accentuels que celles des natifs. La surabondance d'accents dans les productions des apprenants, qui frappent la quasi-totalité des mots lexicaux, a été observé chez des apprenants de diverses L1 (Rasier & Hillingsmann 2007, Kaglik &

Boula de Mareüil 2009, Alazard, Astésano, Billières & Espesser 2011), et peut relever de la non acquisition des règles de désaccentuation du français (Schwab 2012), mais également du propre processus de planification et de formulation en L2 à tous les niveaux linguistiques, entraînant une importante charge cognitive qui se caractériserait par une réduction du débit (*cf. supra*) et par une réduction de l'empan de l'unité de planification, qui se ferait à un niveau plus local que macrostructural (Kaglik & Boula de Mareüil 2009).

Par ailleurs, la probabilité de percevoir 2 groupes accentuels augmente également en l'absence d'information lexicale (séquences délexicalisées *vs* naturelles). Ces résultats confirment l'existence d'une perception différentielle selon si la tâche incite à un traitement de bas niveau ou à un traitement plus proprement linguistique. Ainsi, il semble que, lorsque les sujets entendent des séquences naturelles, ils inhibent plus souvent les marques acoustiques qui leur permettent d'identifier deux groupes accentuels dans les séquences délexicalisées. Néanmoins, cette différence varie entre les deux groupes d'auditeurs : la différence est nettement plus importante pour les francophones natifs que pour les apprenants. Ce phénomène pourrait s'expliquer par le fait que les apprenants traiteraient le signal de parole « naturelle » avec des processus de plus bas niveau (plus acoustiques) tandis que les natifs appliqueraient des processus de plus haut niveau qui inhiberaient les informations acoustiques non pertinentes.

Conformément aux résultats d'études préalables (Padeloup, Espesser & Faraj 2006), nous avons également pu mettre en évidence une relation entre le débit et le nombre de groupes accentuels perçus. Ainsi, plus le débit est lent, plus il est probable que les sujets perçoivent deux groupes au lieu d'un seul. Néanmoins, cette relation diffère en fonction du groupe d'auditeurs. En effet, le débit a un impact nettement plus important sur les natifs que sur les apprenants.

En conclusion, cette étude a montré des différences dans la perception des groupes accentuels entre les francophones natifs et les apprenants hispano-catalanophones de niveau intermédiaire-avancé de français langue étrangère dans des séquences allant au-delà de l'unité lexicale. Ces différences concernent aussi bien le nombre global de groupes accentuels perçus que l'effet de variables telles que le débit ou la présence *vs* absence d'information lexicale, et suggèrent un traitement différentiel de l'information acoustique par les deux types d'auditeurs. Néanmoins, cette étude, encore préliminaire, ne permet pas de conclure si ces effets sont dus aux différences structurelles de l'accent primaire entre le français, d'une part, et l'espagnol et le catalan, de l'autre, ou au propre processus d'apprentissage d'une langue étrangère (Trouvain & Gut 2007).

Bibliographie

- Alazard, C., Astésano, C., Billières, M. & Espesser, R. (2009). Rôle de la prosodie dans la structuration du discours. Proposition d'une méthodologie d'enseignement de l'oral vers l'écrit en Français Langue Étrangère. In Yoo, H.-Y. & Delais-Roussarie, E. (éds) (2011). *Actes d'IDP 2009*, Paris, 9-11 septembre 2009 : 49-61.
- Altmann, H. & Vogel, I. (2002). L2 acquisition of stress: The role of L1. *DGfS Annual Meeting "Multilingualism Today"* in Mannheim, Germany.
- Archibald, J. (1995). The acquisition of stress. In Archibald, J. (éds). *Phonological Acquisition and Phonological Theory*. New Jersey : L. Erlbaum Associates Inc. 81-109.
- Astésano, C., Bertrand, R., Espesser, R. & Nguyen, N. (2012). Perception des frontières et des proéminences en français. *Actes de la conférence conjointe JEP-TALN-RECITAL 2012*, vol 1: 353-360.
- Baayen, R. H., Davidson, D. J. & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items, *Journal of Memory and Language*, 59(4): 390-412.
- Baqué, L. & Estrada, M. (2008). *DESACROM. Corpus pour l'analyse comparée de l'accent dans les langues romanes, version française* [Non publié].
- Barquero Armesto, M. A. (2012). A comparative study on accentual structure between Spanish learners of French interlanguage and French native speakers, *Proceedings of the 6th International Conference on Speech Prosody*, vol. I, pp. 250-253, Shanghai, Chine.
- Boersma, P. & Weenink, D. (2014). Praat: doing phonetics by computer [Logiciel]. Version 5.3.77. [publication électronique]. Consulté sur <http://www.praat.org/>
- Delais-Roussarie, E. & Post, B. (2008). Unités prosodiques et grammaire de l'intonation : vers une nouvelle approche. *Actes des XXVIII^{es} Journées d'Études sur la Parole*, Avignon.
- Dellwo, V. (2012). Sasasa Delexicalizer [Logiciel], version 0.05. [publication électronique]. Consulté sur <http://www.pholab.uzh.ch/static/volker/software.html>
- Di Cristo, A. (1998). Intonation in French. In Hirst, D. & Di Cristo, A. (éds), *Intonation systems: A survey of twenty languages*, Cambridge: Cambridge University Press, 195-218.
- Dresher, E. & Kaye, J. (1990). A computational learning model for metrical phonology. *Cognition*, 34 : 137-195.
- Dupoux, E., Peperkamp, S. & Sebastián-Gallés, N. (2001). A robust method to study stress' deafness. *Journal of the Acoustical Society of America* 110, 1608-1618.

- Estrada, M. (2003). *Contribución al estudio de las manifestaciones fónicas de la afectividad en el habla: la sorpresa como parámetro verbo-tonal*. [Thèse de doctorat non publiée]. Universitat Autònoma de Barcelona.
- Kaglik, A. & Boula de Mareüil, P. (2009). Perception d'un accent étranger et part de la prosodie selon l'âge de première exposition à la L2 : transfert ou phénomène universel en acquisition ? 6^{èmes} Journées d'Etudes Linguistiques, pp. 7-13, Nantes.
- Llisterri, J., Machuca, M.J., de la Mota, C. Riera, M. & Ríos, A. (2005). La percepción del acento léxico en español. In *Filología y lingüística. Estudios ofrecidos a Antonio Quilis* (pp. 271-297). Madrid : CSIC – UNED – Universidad de Valladolid. Vol. 1.
- Martin, P. (2006). Intonation du français : parole spontanée et parole lue. *Estudios de Fonética Experimental*, XV : 133-162.
- Ménétreay, P. & Schwab, S. (sous presse). *Labguistic*: a web platform to design and run speech perception experiments. In Y. Congosto Martín, M. L. Montero Curiel, M.L & A. Salvador Plans (éds): *Fonética Experimental, Espacio Europeo Educación Superior e Investigación*. Universidad de Cáceres y Universidad de Sevilla.
- Padeloup, V., Espesser, R. & Faraj, M. (2006). Sensibilité au débit et marquage accentuel des phonèmes en français. *Actes des XXVI^{es} Journées d'Etudes sur la Parole*, 251-254, Dinard.
- Peperkamp, S. & Dupoux, E. (2002). A typological study of stress 'deafness'. In Gussenhoven, C. & Warner, N. (éds). *Laboratory Phonology*, 7: 203-240. Berlin : Mouton de Gruyter.
- Rasier, L. & Hillingsmann, P. (2007). Prosodic transfer from L1 to L2. Theoretical and methodological issues. *Nouveaux Cahiers de Linguistique Française*, 28 : 41-66.
- Rossi, M. (1999). *L'intonation : le système du français*. Paris : Ophrys.
- Santiago, F. & Delais-Roussarie, E. (2012). What explains the distribution and the form of non-final rising contours in French as an L2? *Proceedings of the International Symposium of Second Language Speech*, 5, 561-579.
- Schwab, S. (2012). Do native Spanish speakers transfer accentual acoustic properties from Spanish to French L2? *Proceedings of the 6th International Conference on Speech Prosody*, vol. I (pp. 330-333), Shanghai, Chine.
- Trouvain, J. & Gut, U. (éds) (2007). *Non-Native Prosody. Phonetic Description and Teaching Practice*. Berlin : Mouton de Gruyter.
- Wunderli, P. (1988). Le débit, indice de l'interrogativité? *Travaux de Linguistique*, 16 : 111-121.