The role of the acoustic correlates of stress in the perception of Spanish accentual contrasts by French speakers

Sandra Schwab [1], *Joaquim Llisterri* [2]

[1] École de Langue et de Civilisation Françaises, Université de Genève;
[2] Departament de Filologia Espanyola, Universitat Autònoma de Barcelona
Sandra.Schwab@unige.ch, Joaquim.Llisterri@uab.cat

Abstract

The aim of this research is to examine the role of the acoustic correlates of lexical stress in the integration of accentual information in French speakers. A shape/pseudoword matching task is used, as it implies not only a low-level acoustic processing, but also a lexical processing. Results show, on the one hand, an influence of the accentual pattern in the perception of stress; on the other, they suggest that French speakers' accentual representations seem to be more rigid than the native Spanish ones.
Index Terms: 'stress deafness', lexical stress contrasts, acoustic correlates, French, Spanish L2.

1. Introduction

French is considered a fixed-stress language, in opposition to free-stress languages, such as Spanish or English. Given that French (primary) stress has a fixed position on the final syllable of the rhythmic group, it can not play a distinctive role at the lexical level, but it rather plays a demarcative function at the utterance level. Therefore, as mentioned by [1], the notion of rhythmic group makes more sense than the notion of word stress in French. As for Spanish, a free-stress language, lexical (primary) stress is realized at the word level. It can appear on one of the last three syllables of the word, which leads to three possible stress patterns: proparoxytone, paroxytone and oxytone. Spanish stress thus plays a distinctive role [2], as it enables the distinction of triplets of words like *número* (['numeɾo], *number*), *numero* ([nu'meɾo], *I number*) and *numeró* ([nume'ɾo], *he/she numbered*). Regarding the acoustic correlates of stress in French and Spanish, both languages use duration, fundamental frequency (F_0) and amplitude in the realization of primary stress. It is well known that French stress is mainly signaled by means of duration, and, to a lesser extent, by F_0 [3, 4], whereas Spanish stress seems to be produced by a combination of duration and F_0 [5].

In view of the accentual differences between French and Spanish, an accentual transfer is likely to occur when French speakers attempt to produce lexical stress in Spanish. The difficulty for French speakers to perceive and produce Spanish stress can be explained by the 'phonological filter' hypothesis [6]. Along the same lines, the notion of 'stress deafness' has been put forward by Dupoux and his coworkers [7, 8, 9]. Using different experimental procedures, they found that sensitivity to stress placement depends on the cognitive load required by the task and on the phonetic variability of the stimuli. Taken together, these experiments lead to the conclusion that French speakers are unable to encode contrastive stress in their phonological representations, although they might be capable, in certain tasks, to make use of the acoustic cues which are present in the speech signal. This might explain the results reported by [10] and [11], showing that French speakers were able to attain relatively high percentages (70% - 83%) of correct identification of the stressed syllable in Spanish words.

As far as the acoustic correlates involved in the perception of stress are concerned, the studies presented in [12] and [13] underline the importance of F_0 variations in Spanish stress identification by French speakers, as it has been shown by [14] for French speakers in French L1. In addition, as shown in [13], the time French speakers need to detect Spanish stress seems to be related to amplitude (alone or combined with duration).

The present research aims at shedding more light on the role of the acoustic correlates of lexical stress (F_0, duration and amplitude) in the integration of the accentual information in French speakers. A shape/pseudoword matching task was used, as it implies not only a low-level acoustic processing but also a lexical processing.

The experiment involved two phases. After a *Training* session, in which participants learned triplets of pseudowords with accentual contrasts by associating these pseudowords to visual shapes (see [15] for a similar experimental design), participants performed, in a *Test* session, the same task on the same pseudowords that they had learned during the training phase, but with acoustic manipulations of F_0, duration, and amplitude.

2. Method

2.1. Participants

Twenty-two native speakers of French (from the French part of Switzerland), aged between 18 and 26 (mean = 20.4), with no knowledge or contact with Spanish or Italian (hereafter 'non-natives', 'NNs'), and 22 bilingual Spanish/Catalan speakers aged between 18 and 31 (mean = 20.7), with no knowledge or contact with French (hereafter 'natives', 'Ns') participated in the study.

2.2. Materials and acoustic stimuli

Two triplets of trisyllabic Spanish pseudowords were taken from the material used in [16]. Each triplet consisted of a proparoxytone (PP; e.g. ***lú**guido*), a paroxytone (P; e.g. *lu**gui**do*) and an oxytone (O; e.g. *lugui**dó***). Following [15], six visual shapes were also created and were randomly pairwise associated with the pseudowords (see Figure 1).

The acoustic stimuli were taken from [16]. In this study, the authors generated *Base* and *Manipulated* stimuli. For this, they created a corpus of 24 trisyllabic words and pseudowords (among them, the 6 used in the present experiment) that a native Spanish speaker read 10 times. For each of the three vowels of each stimulus, the following measures were taken: F_0 at the beginning, at the centre and at the end of the segment; amplitude in five equidistant points along the vowel; and, finally, vowel duration. *Base stimuli* were generated in such a way that the original values of F_0, amplitude and duration were replaced in each vowel of each stimulus by the values averaged over the 10 repetitions. *Manipulated stimuli* were created in the following way: in proparoxytone stimuli, F_0, amplitude and duration values for each vowel were replaced by the corresponding F_0, amplitude and duration values found in the equivalent paroxytone stimuli (henceafter, PP>P stimuli); likewise, in paroxytone stimuli, F_0, amplitude and duration values for each vowel were replaced by the corresponding F_0, amplitude and duration values found in the equivalent oxytone stimuli (henceafter, P>O stimuli). In other words, manipulated stimuli resulted in a shift to the right of the accentual information. The values of each parameter were modified not only individually, but also simultaneously, obtaining seven possible combinations of manipulated parameters; this allows the study of the effects of each acoustic cue both in isolation and in combination with the others. All the manipulations were performed by resynthesis, using the PSOLA algorithm implemented in Praat [17].

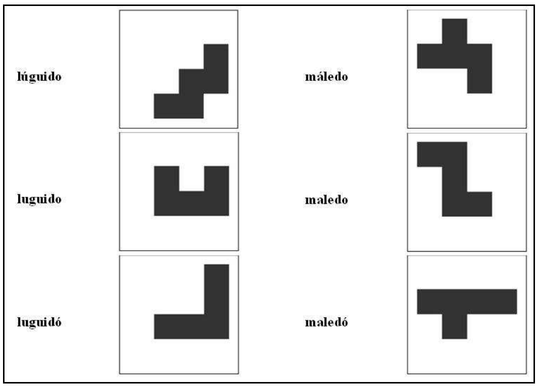

Figure 1: *Pseudowords and visual shapes used in the experiment.*

In the present experiment, the 6 stimuli used in the *Training* session were *Base stimuli*, with no stress shift. In the *Test* session however, the 6 stimuli were *Manipulated stimuli*, i.e. with a stress shift to the right (PP>P and P>O stimuli).

2.3. Procedure

Regarding the *Training* session (see details in [18]), which was divided into 5 blocks, four shapes were presented on the screen and participants heard one of the six pseudowords. They had to click on the shape that they thought corresponded to the pseudoword. In the four first blocks participants received feedback: after each response, the three distractor shapes disappeared, leaving only the correct shape on the screen and the pseudoword was heard again. In the last block, no feedback was given. Only base stimuli were used in this *Training* session, that is, those with no acoustic manipulations. Among the four shapes on the screen, one was the target one (e.g. ***lú****guido*); one was the shape associated with one of the other two members of the triplet (e.g. *lu**gui**do*), and the other two were selected from the three shapes of the other triplet (e.g. ***má****ledo* and *ma**le**do*).

In the *Test* session, we used only manipulated stimuli, which were presented in one block (84 trials). Participants performed the same matching task, but were instructed to answer as quickly as possible and did not receive feedback. Each manipulated stimulus (e.g. PP>P) was presented with the shape corresponding to the original pseudoword (i.e. with stress on the original position; e.g. PP) and with the shape corresponding to the stress-shifted pseudoword (i.e. with the intended shifted stress; e.g. P); the other two shapes were selected from the three shapes of the other triplet (e.g. ***má****ledo* and *ma**le**do*).

Participants were run individually; the stimuli were presented online from a laptop using the DMDX software [19], which recorded the participants' responses and their reaction times.

2.4. Data analysis

Some participants (6 natives and 8 non-natives) had to be excluded because of memorization difficulties and/or an important number of missing data caused by very slow responses.

On the one hand, analyses were performed on the correct/incorrect participants' responses. A *correct* response means that the participant has perceived the stress shift, and thus has clicked on the shape corresponding to the intended position of stress (for example, on the shape of *lu**gui**do* with PP>P manipulated stimuli), and an *incorrect* response means that the participant has not perceived the intended position of the stress, and thus has perceived the original accentual pattern. On the other hand, analyses were performed on reaction times (RTs) in incorrect responses (n=1191), due to the few correct responses for some manipulations. RTs were measured from the end of the stimuli and RTs inferior to 450 ms and superior to 2150 ms were removed (7.53% of the data set).

We analyzed the data (correct/incorrect responses and reaction times) by means of mixed-effects regression models [20], in which participants and pseudowords were entered as random terms. For clarity's sake, the correct/incorrect results and figures are presented in percentages, although all statistical analyses have been performed on raw data.

3. Results and discussion

As results of the *Training* session are fully described in [18], they will only be summarized here. The main finding is that non-natives are able to learn the correspondence between the pseudowords and the shapes: they perform the task reasonably well (73.94% correct responses at the end of the training), although their performance is not as good as the native Spanish one (90.5%) In other words, French speakers are able to learn to perceive lexical stress contrasts after a training session.

Once established that non-natives have the capacity to store and retrieve accentual information, we examined the role of the acoustic correlates of stress (duration, F_0, and amplitude) in the storage of this accentual information. To this end, we considered, for each of the seven manipulations, whether the responses (correct vs. incorrect) and the reaction times are influenced by group (natives and non-natives), and pattern (PP>P, P>O), or by an interaction between these variables. As said above, we studied the impact of the acoustic parameters in the participants' ability to perceive the stress shift (in correct responses), and the eventual time cost produced by the acoustic manipulations when stress shift was not perceived (in the incorrect responses, i.e. when participants perceived the original stress pattern). Note also that, even if the results of various manipulations appear on the same figure (for space reasons), separate analyses were performed for each manipulation.

As far as the isolated manipulation of amplitude (see Figure 2) is concerned, non-natives perceived the stress shift better than natives, whatever the pattern may be (NNs = 16.23%, Ns = 2.86%; $F(1, 328) = 11.15$, $p < .001$). The greater sensitivity to amplitude in non-natives is also observed in reaction times (see Figure 3): when they perceive the original stress pattern, a manipulation of amplitude entails slower reaction times in non-natives than in natives (NNs = 1287.33, Ns = 1169.11; $F(1, 286) = 5.05$, $p < .005$), especially in P>O stimuli ($\beta = 218.10$, $t = 3.01$, $p < .01$).

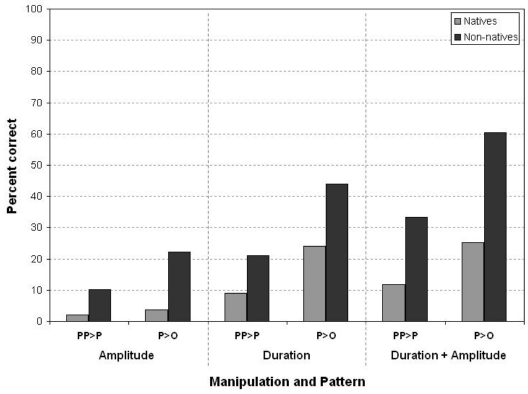

Figure 2: *Percent correct responses as a function of manipulation (Amplitude, Duration, Duration + Amplitude), group and pattern.*

Regarding the isolated manipulation of duration (see Figure 2), non-natives also perceived the stress shift better than natives, whatever the pattern may be (NNs = 35.51%, Ns = 16.56%; $F(1, 315) = 12.78$, $p < .001$). Again, reaction times reflect the greater sensitivity of non-natives than natives (see Figure 3): when they do not perceive the stress shift, non-natives are slower than natives, irrespectively of the pattern (NNs = 1333.70, Ns = 1176.49; $F(1, 229) = 4.32$, $p < .05$).

As for the combined manipulation of duration and amplitude (see Figure 2), non-natives again perceived the stress shift better than natives, whatever the pattern may be (NNs = 46.85%, Ns = 18.54%%; $F(1, 296) = 19.64$, $p < .001$). Nevertheless, when they perceive the original stress pattern, non-natives are not significantly slowed down in their responses (see Figure 3; NNs = 1351.61, Ns = 1254.45; $F(1, 187) = 0.91$, n.s.).

Results for the isolated manipulation of F_0 (see Figure 4) show that non-natives perceived the stress shift better than natives in P>O stimuli (NNs = 46.85%, Ns = 24.79%; $\beta = -1.17$, $z = -2.41$, $p < .05$), whereas natives tend to be better, although not significantly, in perceiving the stress shift in PP>P stimuli (NNs = 30.95%, Ns = 45.73%; $\beta = 0.74$, $z = 1.60$, n.s.). Reaction times show the same trend (see Figure 5): when they perceive the original pattern, non-natives are slower than natives in P>O (NNs = 1344.81, Ns = 1148.50; $\beta = 174.24$, $t = 2.12$, $p < .05$), whereas natives are slower in PP>P (NNs = 1273.17, Ns = 1449.61; $\beta = 190.23$, $t = 2.20$, $p < .05$).

As for the combined manipulation of F_0 and amplitude (see Figure 4), we observe a marginal effect of group in P>O stimuli: non-natives tend to be more sensitive than natives (NNs = 49.63.%, Ns = 35.21%; $\beta = -0.71$, $z = -1.86$, $p = .06$). We note the opposite trend in PP>P stimuli, although the difference is not significant (NNs = 43.21%, Ns = 59.17%; $\beta = 0.64$, $z = 1.65$, n.s.). Reaction times (see Figure 5) show no significant difference between both groups, whatever the pattern may be (NNs = 1374.91, Ns = 1376.14; $F(1,136) = 0.10$, n.s.).

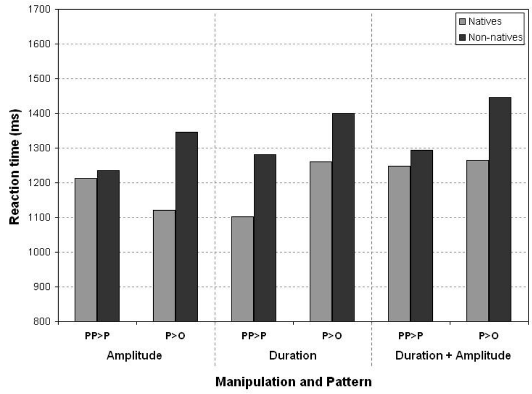

Figure 3: *Reaction times (ms) as a function of manipulation (Amplitude, Duration, Duration + Amplitude), group and pattern.*

As far as the combined manipulation of F_0 and duration (see Figure 4) is concerned, natives perceived the stress shift better than non-natives in P>PP stimuli (NNs = 58.57.%, Ns = 80.10%; $\beta = 1.47$, $z = 2.74$, $p < .01$), while there is no difference between both groups in P>O stimuli (NNs = 67.78.%, Ns = 61.88%; $\beta = -0.03$, $z = -0.07$, n.s). Moreover, whatever the pattern, may be, non-natives, although not significantly, tend to be faster than natives in identifying the original pattern (see Figure 5; NNs = 1219.63.91, Ns = 1400.52; $F(1, 80) = 2.85$, n.s.).

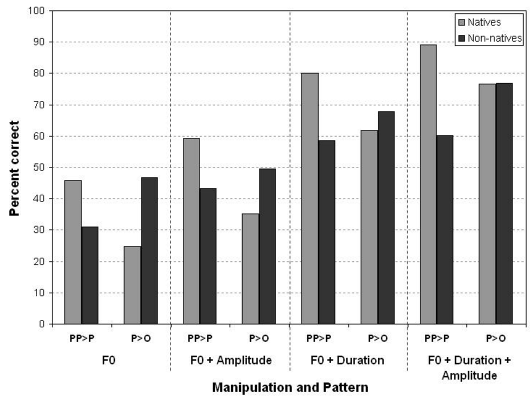

Figure 4: *Percent correct responses as a function of manipulation (F_0, F_0 + Duration, F_0 + Amplitude, F_0 + Duration + Amplitude), group and pattern.*

Finally, when the combined manipulation of the three parameters (see Figure 4) was presented, natives perceived better than non-natives the stress shift in PP>P stimuli (NNs = 60.24%, Ns = 89.17%; $\beta = 2.04$, $z = 3.76$, $p < .001$), while there is no difference between both groups in P>O stimuli (NNs = 76.85%, Ns = 76.67%; $\beta = 0.49$, $z = 0.96$, n.s.). Although natives are slower than non-natives, the difference is not significant (see Figure 5; NNs = 1326.63, Ns = 1454.5; $F(1,54) = 0.46$, n.s.).

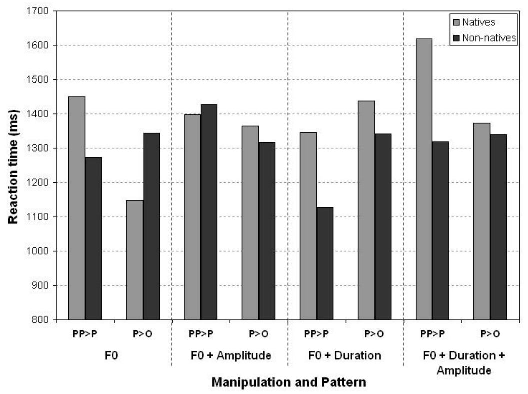

Figure 5: *Reaction times (ms) as a function of manipulation (F_0, F_0 + Duration, F_0 + Amplitude, F_0 + Duration + Amplitude), group and pattern.*

4. Conclusion

Besides showing that French speakers are able to integrate (at least, temporarily) and retrieve the accentual information, results of this research revealed firstly that French speakers do not perceive the stress shift as Spanish speakers do; moreover, these perceptual differences are conditioned by the accentual pattern: while French speakers seem to be more sensitive to an accentual shift to the final syllable, Spanish speakers perceive better a stress shift to the penultimate syllable. Secondly, results highlighted that French speakers, in comparison with Spanish ones, are more sensitive to manipulations of amplitude and duration.

The different behavior of French speakers can possibly be explained by the accentual properties in French, in which lexical stress is generally oxytone and mainly realized by changes in duration [3, 4]. French speakers might have thus transferred this knowledge from French to Spanish. Nevertheless, French speakers are sensitive to amplitude, although it is not an important parameter in the realization of stress in French [3, 4]. This might be a cue to the fact that French speakers are more sensitive than natives to stimuli that present little variations compared with the original ones that they have learned (and stored) during the training session, while native Spanish speakers, used to such variations in their L1, do not perceive them. It might suggest that the accentual representation of the six new Spanish pseudowords acquired and stored in French speakers is more rigid than the representation encoded in Spanish speakers.

Although, these results need to be further examined with stimuli produced by different voices, since phonetic variability in the stimuli has been shown to be important in the study of stress perception [7, 8], this research not only casts some doubt on the existence of 'stress deafness' in French speakers (i.e. they are able to learn to perceive lexical stress contrasts), but also provides promising hints about the integration of accentual information in the L2 lexicon.

5. References

[1] Vaissière, J., "Cross-linguistic prosodic transcription: French vs. English", in N. B. Volskaya, N. D. Svetozarova and P. A. Skrelin [Eds], *Problems and methods of experimental phonetics. In honour of the 70th anniversary of Pr. L.V. Bondarko*, 147-164, St Petersburg: St Petersburg State University Press., 2002.

[2] Quilis, A., *Tratado de fonología y fonética españolas*, Madrid: Gredos, 1993.

[3] Léon, P, *Phonétisme et prononciations du français*, Paris: Armand Colin, 2011.

[4] Léon, P. and Martin, P., "Prosodie et technologie", in E. Guimbretière [Ed], *Apprendre, enseigner, acquérir: la prosodie au cœur du débat*, 135-150, Rouen: Publications de l'Université de Rouen, 2000.

[5] Quilis, A., *Fonética acústica de la lengua española*, Madrid: Gredos, 1981.

[6] Troubetzkoy, N.S., "Grundzüge der phonologie", *Travaux du Cercle Linguistique de Prague* 7, 1939.

[7] Dupoux, E., Pallier, C., Sebastián, N. and Mehler, J., "A destressing 'deafness' in French?", *Journal of Memory and Language* 36(3), 406-421, 1997.

[8] Dupoux, E., Peperkamp, S. and Sebastián, N., "A robust method to study stress 'deafness'", *Journal of the Acoustical Society of America* 110(3-1), 1606-1618, 2001.

[9] Dupoux, E., Sebastián, N., Navarrete, E. and Peperkamp, S., "Persistent stress 'deafness': The case of French learners of Spanish", *Cognition* 106(2), 682-706, 2008.

[10] Muñoz García, M., Panissal, N., Billières, M. and Baqué, L., "¿La metáfora de la criba fonológica se puede aplicar a la percepción del acento léxico español? Estudio experimental con estudiantes francófonos", in Bretones Callejas, M.C., et al. [Eds.], *La lingüística aplicada actual: Comprendiendo el lenguaje y la mente*. Almería: Universidad de Almería - AESLA, Asociación Española de Lingüística Aplicada, 489-499, 2009.

[11] Schwab, S. and Llisterri, J., "La perception de l'accent lexical espagnol par des apprenants francophones", in Baqué, L. and Estrada, M. [Eds.], *La langue et l'être communiquant. Hommage à Julio Murillo*. Mons: Éditions du CIPA, 311-328, 2010.

[12] Mora, E., Courtois, F. and Cavé, C., "Étude comparative de la perception par des sujets francophones et hispanophones de l'accent lexical en espagnol", *Revue Parole* 1, 75-86, 1997.

[13] Schwab, S. and Llisterri, J., "The perception of Spanish lexical stress by French speakers: Stress identification and time cost", in Wrembel, M., Kul, M., Dziubalska-Kołacyk, K., [Eds.], *Achievements and perspectives in SLA of speech: New sounds 2010*. Vol. 1. Frankfurt am Main: Peter Lang, 229-242, 2011.

[14] Rigault A., "Rôle de la fréquence, de l'intensité et de la durée vocalique dans la perception de l'accent en français", in *Proceedings of the 4th International Congress of Phonetic Sciences*, The Hague: Mouton, 735-748, 1962.

[15] Dufour, S., Nguyen, N. and Frauenfelder, U., "Does training on a phonemic contrast absent in the listener's dialect influence word recognition?", *Journal of the Acoustical Society of America* Express Letters 128, EL43-EL48, 2010.

[16] Llisterri, J., Machuca, M.J., de la Mota, C., Riera, M. and Ríos, A., "La percepción del acento léxico en español", in *Filología y lingüística. Estudios ofrecidos a Antonio Quilis*. Madrid: Consejo Superior de Investigaciones Científicas – Universidad Nacional de Educación a Distancia – Universidad de Valladolid, 271-97, 2005.

[17] Boersma, P. and Weenink, D., Praat: doing phonetics by computer (Version 5.2). www.praat.org, 2011.

[18] Schwab, S. and Llisterri, J., "Are French speakers able to learn to perceive lexical stress contrasts?", in *Proceedings of the 17th International Congress of Phonetic Sciences*. Hong Kong, China. 17-21 August, 2011, 1774-1777.

[19] Forster, J. C., DMDX updates page. Online: http://www.u.arizona.edu/~jforster/dmdx.htm, accessed on Nov 4 2011.

[20] Baayen, R.H., Davidson, D.J. and Bates, D.M., "Mixed effects modeling with crossed random effects for subjects and items", *Journal of Memory and Language* 59, 390-412, 2008.