TO WHAT EXTENT DO FRENCH LISTENERS PERCEIVE DIFFERENCES IN ACCENT LOCATION WITHIN WORDS? AN EEG INVESTIGATION

Outhmane Rassili, Amandine Michelas, Sophie Dufour

Aix Marseille Université, CNRS, LPL, Aix-en-Provence
outhmane.rassili@univ-amu.fr, amandine.michelas@univ-amu.fr, sophie.dufour@univ-amu.fr

ABSTRACT

In this study, we provided a more in-depth examination of the ability of French listeners to perceive and use the location of accent in a discrimination task. Event-related potentials (ERPs) were recorded while participants performed a same-different task. Different stimuli diverged either in one phoneme (e.g., /'**ʒy**ʁi/-/'**ʒy**ʁɔ̃/) or in accent location (e.g., /'**ʒy**ʁi/-/ʒy'**ʁi**/). Although, participants reached 90% of correct responses, ERPs results indicated that a change in accent location was not detected while a change in one phoneme was detected in two time-windows, the first between 400 and 500 ms, and the second between 500 and 650 ms after target onset. It results that in the early moments of speech processing, stimuli that are phonemically identical but that differ in accent location are perceived as being strictly similar. The good performance observed in behavioral responses was interpreted in relation to attentional processes.

1. INTRODUCTION

Is /'**bo**pelo/ different from /bo'**pe**lo/? The answer is undoubtedly positive for listeners of lexical-stress languages such as Spanish for which accent location is lexically contrastive. However, what about for listeners of languages such as French for which accent doesn't not change the meaning of words[1]? In a well-known study, Dupoux et al. [1] have compared the performance of French and Spanish listeners in the discrimination of accentual patterns. They reported that French listeners had more difficulties than Spanish listeners in discriminating between non-words that differed in accent location ('**bo**pelo, bo'**pe**lo). They also reported that French listeners had more difficulties in discriminating between non-words that differed in accent location ('**bo**pelo, bo'**pe**lo) than between non-words that differed in one phoneme ('**bo**pelo,'**so**pelo), while Spanish listeners performed equally well on accentual and phonemic contrasts. The French listeners difficulties in perceiving accent location has been replicated in several subsequent studies when tested on non-words [e.g., 2-6] and both on Spanish [7] and French words [8]. Although the difficulties of French listeners in distinguishing stimuli that differ in accent location are manifest when their performance is compared to that of Spanish listeners, and perhaps more crucially when their performance is compared to their own performance on phonemic contrasts, it should be noted that in all of the studies, French listeners performed rather well, above the chance level (around 85% of correct responses; [see in 1,8]). Hence, even if their performance was not as high as that of listeners of languages with a lexical stress, and not as high as their performance on phonemic contrasts, French listeners showed a certain capacity to perceive and use accent location to discriminate between stimuli that are phonemically identical.

In this study, we pursued these investigations and provided a more in-depth examination of the ability of French listeners to perceive and use accent location in a discrimination task. More particularly, we conducted an electroencephalographic (EEG) study known to reflect cognitive processes as it unfolds over time[2]. This allowed us to examine the earliest moment at which French listeners perceive the difference in accent location, in contrast to classic behavioral tasks that allow examining only the final product of perception. We thus recorded ERPs while

[1] In French, a word can be accented or not depending on its location within the accentual phrase. For instance, the word *ballon* "ball" receives accent in ([*un petit ba'**LLON***]) "a little ball" because it is in final position within the phrase while it is unaccented in ([*un baLLON '**bleu***]) "a blue ball" because it is not in final position. Note also that the term " accent" was used each time we refer to the acoustic prominence of a syllable.

[2] Although a number of studies have used EEG to examine the representation and perception of accent in listeners of lexical-stress languages (e.g., [9-12]), to the best of our knowledge, only one EEG study [13] has been conducted to examine the French listeners' perception of differences in accent location. However, in this study, EEG measurements were conducted after a training session (with no EEG session before) in which participants learned to distinguish between tri-syllabic stimuli differing in accent location. Here, French listeners' perception was examined without implicit or explicit training which allowed us to examine "basic" perception.

French participants performed a same-different task. They first heard four stimuli that were strictly identical in respect to both their phonemic and accentual patterns, but produced by four different female speakers (e.g., /**'ʒy**ʁi/-/**'ʒy**ʁi /-/**'ʒy**ʁi /-/**'ʒy**ʁi/), and then heard a fifth stimulus, the target, always produced by a male speaker (see [14] for the same design with ERPs). The target was either the same as (identical condition; e.g., /**'ʒy**ʁi/), or different from the first four stimuli (deviant conditions). In the deviant conditions, the target stimulus differed from the first four stimuli, either in the phonemic (phonemic deviant condition; e.g., /**'ʒy**ʁɔ̃/) or in the accentual (accentual deviant condition; e.g., /ʒy**'ʁi/**) pattern. We measured the precise moment(s) at which phonemic and accentual discrimination occurs by comparing the ERPs in the standard vs. deviant conditions.

2. METHODOLOGY

2.1. Participants

20 right-handed French native speakers (15 females) between 18 and 33 years old (mean=22.4, SD=3.6) participated in the experiment. Since English (or German) are taught from middle school in the French educational system, they all reported having knowledge of one of these languages (18 in English and 2 in German). Some of them also reported having knowledge in Italian, Spanish, Russian, Portuguese or Mandarin Chinese, but none spent more than 3 months abroad, and all have been exposed to these languages after 10 years of age during middle school. No participant reported having any neurological, hearing or language impairment. Each participant gave written informed consent prior to the experiment, and they received either a course credit or 20€ for their participation. One participant was removed due to a high error rate (>50%).

2.2. Materials

Two bisyllabic French words *jury* /ʒyʁi/ “jury” and *juron* /ʒyʁɔ̃/ “swearword”, differing only on the last phoneme were selected. Five French speakers (four women, one man) produced the two words within carrier sentences of two types. In the first type, target words were accented on the last syllable (e.g., *[On m’avait parlé] [d’un exigeant JU'**RY**] [qu’il était difficile] [de satisfaire]* “I was told about a highly demanding jury that it was difficult to satisfy”. Because words bearing primary accent on the first syllable are not found in French, we used sentences in which target words are unaccented (e.g., *[On m’avait parlé] [d’un JURY exi'**geant**] [qu’il était difficile] [de satisfaire]*) but the speakers were explicitly instructed to produce the target words by making the first syllable prominent. To avoid coarticulation effects due to contextualized-speech, the two versions of each word (accented on the first syllable, accented on the second syllable) were then extracted from the carrier sentences and auditorily presented in isolation to each speaker. The speakers were instructed to imitate each version of the words. The recordings were made in a sound-attenuated room, and digitized at a sampling rate of 48,000 kHz. Each word in each of the two versions was then normalized at a level of 60 dB. Acoustic analyses were conducted to ensure that the words were produced with the expected accentual patterns. The length of the two syllables and the f0 movement associated with each syllable were measured (see Table 1).

Speaker	Word accented on the	1st syllable				2nd syllable			
		Syll. dur. (ms)	(Hz)	(Hz)	Pitch slope (%)	Syll. Dur. (ms)	(Hz)	(Hz)	Pitch slope (%)
Male speaker	1st syll.	264	115	153	33	176	98	85	-13
	2nd syll.	163	99	94	-5	275	100	244	146
Female speakers	1st syll.	343	214	304	42	179	224	208	-7
	2nd syll.	182	193	191	-1	286	187	447	141

Table 1. Acoustic properties of the target words in their two accentual patterns for the five speakers. For the 4 female speakers, means are reported.

2.3. Procedure

Stimuli were presented binaurally via headphones in an acoustically shielded room. Trials consisted in five stimuli each separated from the following one by 600 ms of silence. On any given trial, the first four stimuli were strictly identical in respect to both their phonemic and accentual patterns, but produced by the four female speakers. The order of the female speakers was counterbalanced across the trials. The fifth and final stimulus, the target stimulus, always produced by the male speaker, was either the same as (identical condition), or different from the first four stimuli (deviant conditions). In the deviant conditions, the target stimulus differed from the first four stimuli, either in the phonemic (phonemic deviant condition) or in the accentual (accentual deviant condition) pattern. The experimental conditions are illustrated in Table 2.

Condition	Female speakers	Male speaker	Number of trials	Expected response
Identical	/ʒy'**ʁi**/	/ʒy'**ʁi**/	48	Same
	/'**ʒy**ʁi/	/'**ʒy**ʁi/	48	Same

Accentual	/**'ʒy**ʁi/	/ʒy**'ʁi**/	48	Different
dev.	/ʒy**'ʁi**/	/**'ʒy**ʁi/	48	Different
Phonemic	/ʒy'**ʁɔ̃**/	/ʒy'**ʁi**/	48	Different
dev.	/**'ʒy**ʁɔ̃/	/**'ʒy**ʁi/	48	Different

Table 2: Illustration of the experimental conditions with the *jury* stimuli.

96 trials were presented per condition. So that there was an equal number of trials requiring a "same" and "different" response, 96 filler trials constructed exactly on the same model as the trials used in the identical condition were added, but were not included in the analyses. Each participant heard a total of 384 trials that were presented randomly. They were instructed to respond same when the target word was pronounced both with the same sounds and the same accentuation as the four preceding words, and to respond different when the target word differed from the preceding words either in its sounds or in its accentuation. Different responses were made using their dominant hand on a response box that was placed in front of them. Reaction Times (RTs) were measured from the onset of target stimuli. An inter-trial interval of 3000 ms elapsed between the end of one trial and the beginning of the next. Participants first received 6 practice trials and then three blocks of 128 trials separated by a break. No feedback was given to participants. The experiment lasted 1 h.

3. DATA ACQUISITION & ANALYSIS

The EEG signal was recorded from the scalp with a 64-channel BioSemi Active-Two AD-box. Individual electrodes were adjusted to a stable offset lower than 20 mV, and the EEG signal was digitized at 512 Hz. The EEG epochs, starting at 100 ms before the test stimulus and ending 700 ms after it, were averaged for each experimental condition and for each participant. The EEG data were filtered offline by a bandpass (1–30 Hz), re-referenced offline to the average of left and right mastoids, and corrected by a baseline of 100 ms before the onset of the target stimulus. Epochs were accepted under an artifact rejection criterion of 100 μV. All participants had a number of accepted trials superior to 80 for each experimental condition. For each participant, data from bad channels were interpolated.

An initial inspection of ERP waveforms showed differences between the three conditions at frontocentral sites, and on two negative components. Two topographical sites of 9 electrodes were used for the analyses: left frontocentral (F5, F3, F1, FC5, FC3, FC1, C5, C3, C1) and right frontocentral (F6, F4, F2, FC6, FC4, FC2, C6, C4, C2). To assess the timing of differential ERPs between the conditions, we conducted an ANOVA with the factors Condition (identical, phonemic deviant, accentual deviant) and Hemisphere (left, right). Two time-windows were selected around the peak amplitude of the two negativities identified: 400-500 ms and 500-650 ms. The Greenhouse-Geisser correction was applied and the corrected *p*values are reported.

4. BEHAVIOURAL RESULTS

ANOVAs were also performed on RTs and Error rates for the "different responses". The RT analysis was performed on the correct responses. An inspection of the data revealed that two short RTs (<300 ms) and three extremely long RTs (>10,000 ms) strongly deviated from the distribution, and were thus discarded from the analysis. A significant difference was found between the two deviant conditions on both RTs ($F(1,18)=9.95$, $p<.01$) and Error rates ($F(1,18)=4.44$, $p<.05$). Participants responded slower and with more errors in the accentual (Mean RTs=1158 ms; Mean Error rates=6.80%) than in the phonemic deviant condition (Mean RTs=1044 ms; Mean Error rates =1.32%).

5. ELECTROPHYSIOLOGICAL RESULTS

A main effect of the Condition was observed in the two time-windows with more negative values in the phonemic deviant condition than in both the accentual deviant and the identical conditions (see Table 3). No difference was observed between the accentual deviant and the identical condition. The interaction between Condition and Hemisphere was also significant, and was due to a greater effect of Condition in the right hemisphere in the first time-window, and to a greater effect of Condition in the left hemisphere in the second time-window. Figure 1 displays grand-average waveforms in each condition at FC1 electrode.

	400-500ms	500-650ms
Effect of Condition	**p<.001**	**p<.01**
Identical vs. Phon deviant	p<.01	p<.001
Identical vs. Acc deviant	p>.20	p>.20
Phon deviant vs. Acc deviant	p<.01	p<.01
Hemisphere*Condition	**p<.05**	**p<.05**
Left Hemisphere		
Identical vs. Phon deviant	p<.01	p<.001
Identical vs. Acc deviant	p=0.14	p>.20
Phon deviant vs. Acc deviant	p<.05	p<.05
Right Hemisphere		
Identical vs. Phon deviant	p<.01	p<.01
Identical vs. Acc deviant	p>.20	p>.20
Phon deviant vs. Acc deviant	p<.01	p<.01

Table 3. Summary of the ANOVA results.

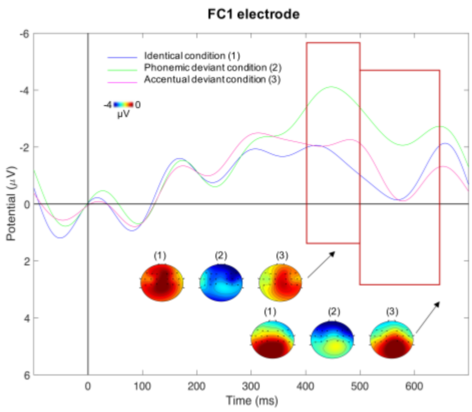

Figure 1. Grand-average waveforms at FC1 electrode in each condition. The maps show the topography of the grand-averaged ERPs for each condition at 400–500ms and 500–650ms.

6. DISCUSSION

ERP results clearly indicated that stimuli that involve a change in accent location are processed differently than stimuli that involve a change in one phoneme. In particular, we observed that a change in one phoneme was detected in two time-windows, respectively between 400-500 ms and between 500-650 ms after target onset, while a change in accent location was not detected. As a result, in the early moments of speech processing, stimuli that are phonemically identical but that differ in accent location are perceived as being strictly similar, and thus no difference between the accentual deviant and the identical conditions has been found. Because in French, accent location contrasts do not exist, our results are in perfect line with all of studies that report that listeners do not detect a phonemic change when the phonemic contrast under investigation does not exist in the native language [14]. For example, using the same design as in the present study, [14] showed that French listeners did not perceive the difference between the dental /da/ and the retroflex /Da/ that do not exist in French. Hence, it appears that when a contrast is not present in the native language, listeners fail to perceive it, at least when we probe processing in the early moments, and this seems to be true for both accentual and phonemic contrasts.

It remains however to explain why our participants succeeded in doing the task with a level of performance that reached 93% for the accentual contrast. As participants were asked to judge whether the target was identical or different from the previous stimuli, we believe that this good performance reflects only decisional processes linked to this particular task, and not unconscious automatic speech perception processes. Moreover, the two types of differences present in the stimuli were explained to the participants during the instructions, and thus we have likely attracted the participants' attention on the differences. Hence, if our reasoning is correct, French listeners would be able to use accent location in order to perform some tasks provided that their attention is directed on the accentual differences. The same conclusion can be drawn from a recent study [13] who examined the ability of French listeners to discriminate accent location contrasts after training. In particular, using ERPs measurement in an active deviant detection task, the authors reported evidence for discrimination on the P3 component known to reflect conscious and attentional processes. However, no evidence was found earlier in the processing, before the P3, and on the N2b component known to reflect automatic and unconscious processes. Further research should thus be done to compare French listeners performance in tasks that attract participant's attention on the differences under investigation, and in tasks that disengage participant's attention from these differences. Despite the good discrimination at a decisional level, behavioral measurements nonetheless revealed more errors and longer RTs in the accentual than in the phonemic deviant condition. This confirms the greater difficulty that French listeners have to use accent location information in comparison to phonemic information in order to perform a discrimination task, likely because it is not usual for French listeners to process accent location.

To conclude, despite a high level of performance in discrimination tasks, French listeners do not perceive the difference between stimuli like /**'ʒy**ʁi/ and /ʒy**'ʁi/** during unconscious and automatic speech perception processes. Although discrimination tasks provide important insights regarding conscious and decisional processes, they nonetheless tend to overestimate the ability to perceive and use accentual contrasts. Hence, it seems crucial to test French listeners' ability to perceive and use accent in a wide variety of tasks with experimental paradigms enable to probe both unconscious and conscious processes.

7. REFERENCES

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

[2] Dupoux, E., Peperkamp, S., and Sebastián-Gallés, N. 2001. A robust method to study stress 'deafness', *J. Acoust. Soc.* 110(3), 1606–1618.

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

[4] Peperkamp, S., Vendelin, I., Dupoux, E. 2010. Perception of predictable stress: A cross-linguistic investigation. *J. of Phonetics 38*(3), 422-430.

[5] Rahmani, H., Rietveld, T., Gussenhoven, C. 2015. Stress “deafness” reveals absence of lexical marking of stress or tone in the adult grammar. *PloS one* 10(12), e0143968.
[6] Carpenter, A. 2015. Phonetic training significantly mitigates the stress 'deafness' of French speakers. *Intern. J. of Linguistics* 7.
[7] Schwab, S., Llisterri, J. 2013. Does training make French speakers more able to identify lexical stress?. *Proceedings of the international symposium on the acquisition of second language speech, Concordia working papers in applied linguistics:* 5.
[8] Michelas, A., Esteve-Gibert, N., Dufour, S. 2018. On French listeners’ ability to use stress during spoken word processing. *J. of Cognitive Psychology 30*(2), 198-206.
[9] Domahs, U., Wiese, R., Bornkessel-Schlesewsky, I., & Schlesewsky, M. 2008. The processing of German word stress: evidence for the prosodic hierarchy. *Phonology* 25(1), 1-36.
[10] Molczanow, J., Domahs, U., Knaus, J., & Wiese, R. 2013. The lexical representation of word stress in Russian: Evidence from event-related potentials. *The mental lexicon* 8(2), 164-194.
[11] Zora, H., Schwarz, I.-C., & Heldner, M. 2015. Neural correlates of lexical stress: Mismatch negativity reflects fundamental frequency and intensity. *NeuroReport* 26(13), 791–796.
[12] Lu, S., Correia, S., Jeronimo, R., Vigário, M., & Frota, S. 2016. Revisiting stress 'deafness' in European Portuguese - an ERP study. In J. Barnes, S. Shattuck-Hufnagel, N. Veilleux, & Orgs (Eds.). *Proceedings of speech prosody 2016 Online publication.* ISSNhttps://doi.org/10.21437/SpeechProsody.2016 2333-2042.
[13] Schwab, S., Giroud, N., Meyer, M., & Dellwo, V. 2020. Working memory and not acoustic sensitivity is related to stress processing ability in a foreign language: An ERP study. *J. of Neurolinguistics 55*, 100897.
[14] Dehaene-Lambertz, G. 1997. Electrophysiological correlates of categorical phoneme perception in adults. *NeuroReport* 8(4), 919-924.