



**HAL**  
open science

## How deaf are French speakers to stress?

Amandine Michelas, Ulrich Hans Frauenfelder, Daniele Schoen, Sophie Dufour

► **To cite this version:**

Amandine Michelas, Ulrich Hans Frauenfelder, Daniele Schoen, Sophie Dufour. How deaf are French speakers to stress?. *Journal of the Acoustical Society of America*, 2016, 139 (3), pp.1333-1342. 10.1121/1.4944574 . hal-01431279

**HAL Id: hal-01431279**

**<https://hal.science/hal-01431279>**

Submitted on 23 Jul 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Title:** How deaf are French speakers to stress?

**Running title:** Stress in French word perception

**Authors:**

Amandine Michelas <sup>a)</sup>

Aix Marseille Université, CNRS, LPL, UMR 7309, 5 Avenue Pasteur, 13100 Aix-en-Provence, France

Ulrich Hans Frauenfelder

Laboratoire de Psycholinguistique Expérimentale, FPSE, Université de Genève, Boulevard du Pont d'Arve 40, 1204 Genève, Switzerland

Daniele Schön<sup>b)</sup>

Aix-Marseille Université, INS, 27 Boulevard Jean Moulin, 13005 Marseille, France

Sophie Dufour

Aix Marseille Université, CNRS, LPL, UMR 7309, 5 Avenue Pasteur, 13100 Aix-en-Provence, France

<sup>a)</sup> Electronic mail : [michelas@lpl-aix.fr](mailto:michelas@lpl-aix.fr)

<sup>b)</sup> Also at : INSERM, U1106, 27 Boulevard Jean Moulin, 13005 Marseille, France

## **Abstract**

This ERP study examined whether French listeners use stress at a phonological level when discriminating between stressed and unstressed words in their language. Participants heard five words and made same/different decisions about the final word (male voice) with respect to the four preceding words (different female voices). Compared to the first four context words, the target word was (i) phonemically and prosodically identical (/fu/-/fu/; control condition), (ii) phonemically identical but differing in the presence of a primary stress (/fu'/-/fu/), (iii) prosodically identical but phonemically different (/fo/-/fu/), or (iv) both phonemically and prosodically different (/fo'/-/fu/). Crucially, differences on the P200 and the following N200 components were observed for the /fu'/-/fu/ and the /fo/-/fu/ conditions compared to the /fu/-/fu/ control condition. Moreover, on the N200 component more negativity was observed for the /fo/-/fu/ compared to the /fu'/-/fu/ conditions, while no difference emerged between these two conditions on the earlier P200 component. Crucially, our results suggest that French listeners are capable of creating an abstract representation of stress. However, as they receive more input, participants react more strongly to phonemic than to stress information.

## I. INTRODUCTION

Perhaps one of the most striking characteristics of French prosody compared to that of other Romance languages is that “stress” (or “accent”, i.e., the phonetic/perceptual prominence of a syllable) does not carry lexical information and is a property of a larger unit than the word. In Spanish, in contrast, stress affects the lexical level and can change the meaning of a word, as in the case of */be’be/* ‘s/he drinks’ and */bebe’/* ‘baby’. In French, primary stress always affects the last syllable of a larger unit, that is, the accentual group. Although there is also an optional secondary stress in French which is associated with the beginning of the accentual group, we will focus here on primary stress. The fact that primary stress is always final within the accentual group has at least two main consequences. First, it is not possible to find minimal pairs of words which differ only in stress position. Second, a same word receives primary stress or not depending on its position within the accentual group. For instance, the monosyllabic word *choux* ‘cabbage’ receives primary stress in the following sentence [*Le livre*] [*racontait l’histoire*] [*d’un petit CHOU*] [*qui était mangé*] [*par une sorcière*] ‘The book told the story of a little cabbage that was eaten by a witch’ because it is the last full syllable of the accentual phrase. In contrast, it is unstressed in the sentence [*Le livre*] [*racontait l’histoire*] [*d’un CHOU bleu*] [*qui souffrait*] [*de sa différence*] ‘The book told the story of a blue cabbage suffering from its difference’ because it is not in final position within the accentual phrase.

Studies using on-line measures of spoken word recognition have shown that stress in French can help listeners to resolve lexical ambiguity at the phrasal level. In their seminal study, Christophe, Peperkamp, Pallier, Block & Mehler (2004) asked participants to detect a target word (e.g., *CHAT*) in sentences like [*Le CHAT grincheux*] ‘the grumpy cat’ or [*Le CHAT*] [*grimpaît*] ‘the cat climbed up’. These authors found slower detection times in [*Le CHAT*

*grincheux*'] than in [*Le CHAT*] [*grimpait*']. The slower detection of *CHAT* in [*Le CHAT grincheux*'] has been interpreted as resulting from competition between *CHAT* 'cat' and its competitor *CHAGRIN* 'sadness'. In this case, since there is no stress on *CHAT* to indicate its end, the word *CHAGRIN* remains active and competes with the word *CHAT*. In the same vein, Spinelli and collaborators (Spinelli, Welby & Schaegis, 2007; Spinelli, Grimault, Meunier & Welby, 2010) showed that French listeners are able to use optional secondary stress on the first syllable of the first content word of the accentual group to differentiate between phonemically ambiguous sequences such as *l'Affiche* (/l#a#fiʃ/, 'the poster') vs. *la Fiche* (/la#fiʃ/) 'the sheet').

Although French listeners use stress to locate the end of stressed words (Rietveld, 1980; Bannel & Bacri, 1994; Christophe et al., 2004), they experience difficulties in distinguishing between nonsense words that differ only in stress position (Dupoux, Pallier, Sebastian & Mehler, 1997; Dupoux, Peperkamp & Sebastian-Galles, 2001; Dupoux, Peperkamp & Sebastian-Galles, 2010; Peperkamp, Vendelin & Dupoux, 2010). In an ABX task, in which A, B and X were spoken by three different speakers, and in which participants judged whether X was identical to A or B, Dupoux et al. (1997) reported that French listeners had more difficulties in discriminating between two non-words that differed only in stress position (*fi' dape*, *fida' pe*, *fi' dape*) than when the two non-words differed only in their first phoneme (*fi' dape*, *li' dape*, *fi' dape*). French listeners also had more difficulties than Spanish listeners in distinguishing two non-words that differed only in stress position. Further, Spanish listeners, unlike French listeners, performed equally well on stress and phoneme contrasts and benefited additionally from stress information, in a condition where phonemic information was sufficient to perform the task (*fi' dape*, *lida' pe*, *fi' dape*). It was only when the stimuli in a standard AX task were produced by a single speaker that French listeners performed equally well on stress (*fi' dape* vs.

*fi'da'pe*) and phoneme (*fi'dape* vs. *li'dape*) contrasts (see also Dupoux et al., 2001 with a sequence repetition task).

The studies of Dupoux et al. (1997, 2001) thus revealed that French listeners are quite capable of using an acoustic representation of stress to discriminate between two items that are pronounced by the same speaker. Importantly, these studies also showed that performance differences between stress and phoneme contrasts only emerged when acoustically-based responses were made impossible by the use of different speakers, even if the French participants in the study of Dupoux et al. (1997) still performed much better than chance (50%) on the stress contrast. Hence, French listeners appear to have difficulties in using stress when they are forced to encode it phonologically, and thus represent it more abstractly. However, this difficulty of French participants could perhaps be explained by these authors' use of stimuli that do not respect the rules of stress placement in French. In this study 1st vs. 2nd stressed syllables (*bo'pelo* vs *bop'e'lo*) and 2nd vs. 3rd stressed syllables (*bop'e'lo* vs *bopelo'*) were used, but in French, a word cannot receive primary stress at any non-final position.

In this study, we have thus re-examined the ability of French listeners to use stress information at a phonological level of processing with stimuli that respect the rules of stress placement in French. Our stimuli consisted of common stress patterns, namely monosyllabic target words that were unstressed or bore primary stress (e.g., *chou* vs. *chou'*). Our target words were extracted from carrier sentences (see method) in which the target word bore primary stress or was unstressed depending on its position within the utterance (e.g., [*Le livre*] [*racontait l'histoire*] [*d'un petit CHOU*] [*qui était mangé*] [*par une sorcière*] vs. [*Le livre*] [*racontait l'histoire*] [*d'un CHOU bleu*] [*qui souffrait*] [*de sa différence*]). Hence, with this particular stress contrast, we tested French listeners' ability to discriminate between stressed and

unstressed words rather than their ability to locate stress position within the word. As in Dupoux et al.'s study, a phonemic contrast condition involving two words differing in the last phoneme (e.g., **chou** 'cabbage' vs. **chaud** 'hot') was used. This allowed us to examine the relative contributions of stress and phonemic information in French word perception. Finally we also tested whether French listeners benefit from stress information when phonemic information alone is sufficient to differentiate the two members of a minimal pair (e.g., **chou** vs. **chaud**).

To collect detailed information on the perceptual processing of stress and phonemic information over time, we recorded event-related potentials (ERP) while French participants performed a same-different task. They first heard four different pronunciations of the same word by 4 female speakers (e.g., /ʃu/-/ʃu/-/ʃu/-/ʃu/), and then heard the target word produced by a male speaker. With respect to the context tokens, this target could be either (i) phonemically and prosodically identical (/ʃu/; *phon&pros identical* condition), (ii) phonemically identical but differing from the preceding word in the presence of a primary stress (/ʃu'/; *pros dev* condition), (iii) prosodically identical but phonemically deviant (/ʃo/; *phon dev* condition), or (iv) both phonemically and prosodically deviant (/ʃo'/; *phon&pros dev* condition). Acoustic variability was introduced into the stimuli by mixing the productions of different speakers in order to prevent participants from basing their decision only on acoustic changes, and to force them to rely on a more abstract level of representations (see also Brunellière, Dufour, Nguyen & Frauenfelder, 2009; Dehaene-Lambertz, Dupoux & Gout, 2000; Dupoux et al., 2001; Eulitz & Lahiri, 2004). We focused on two ERPs components typically observed in this kind of paradigm. The first is a positive component peaking between about 150 and 275 ms after the onset of the stimulus (P200). The second is a frontal negative component (N200) peaking between 250 and 350 ms after the stimulus onset (Dehaene-Lambertz, 1997; Dehaene-Lambertz et al., 2000; Brunellière et al., 2009). Both the P200 and

N200 components are known to reflect phonological processes and have been shown to be sensitive to phonological changes (Dehaene-Lambertz, 1997; Dehaene-Lambertz et al., 2000; Brunellière et al., 2009). Note that the experimental procedure used in our study differs from that used in the classical oddball paradigm in which the so-called mismatch negativity (MMN) component is observed (Näätänen, Simpson & Loveless, 1982, Näätänen et al., 1997). In particular, our participants were engaged in an active task, whereas participants in the classic MMN procedure are engaged in a passive listening paradigm requiring no response (e.g., Näätänen et al., 2007).

First, we hypothesized that if French participants exploit stress information at a phonological level, amplitude differences on both the P200 and N200 components should be observed in the *pros dev* (e.g. /**ju**'-/ /ju/) compared to the *pros&phon identical* (e.g. /ju-/ /ju/) condition. We also hypothesized that if French listeners use primary stress differences in the same way as they use phonemic differences to distinguish between two words, both the *pros dev* (e.g. /**ju**'-/ /ju/) and *phon dev* (e.g. /fo-/ /ju/) conditions should exhibit P200 and N200 amplitude differences compared to the *pros&phon identical* (e.g. /ju-/ /ju/) condition, while no difference should be observed between the two first conditions (/**ju**'-/ /ju/ and /fo-/ /ju/ conditions). Finally, if French participants use stress information even when phonemic information is sufficient to distinguish between the target word and the preceding word, differences should also be observed between the *phon dev* (/fo-/ /ju/) condition and the *pros&phon dev* condition (/fo' /- /ju/). Note that differences on a third component in a later time window, namely the LPC (late positive component) which is more related to conscious phonological decision processes (Dehaene-Lambertz et al., 2000; Brunellière et al., 2009) will also be examined.



## II. MATERIALS AND METHOD

### A. Participants

19 participants (10 women) between 18 and 36 years old (mean = 21 years, 5 months) participated in the experiment. All were right-handed French-speaking students from Aix-Marseille University. Before the experiment, participants were asked to complete a questionnaire including information about any hearing problems and/or language impairments. They all reported having no neurological/hearing impairment or language disorders. All the participants were native speakers of French. Among them, 15 spoke English; 4 judged their English proficiency as poor, 6 as medium, 4 as good and 1 as excellent. 3 of the 19 participants spoke Italian (one with excellent proficiency and two as beginners). Three of them were also beginners in Spanish, one of them in German and one in Chinese. Only one participant spent more than three months abroad, and all have been exposed to a second language after 10 years of age during middle school. Each participant gave informed consent prior to experimentation and was remunerated for their time with 15 euros. EEG data from two participants were excluded due to a very poor signal.

### B. Stimuli

Five French speakers (four women, one man) produced the two French words /ʃu/ (chou ‘cabbage’), and /ʃo/ (chaud ‘hot’) within carrier sentences in which the target word bore primary stress or was unstressed depending on its position within the utterance (e.g., [*Le livre*] [*racontait l’histoire*] [*d’un petit **CHOU***] [*qui était mangé*] [*par une sorcière*] vs. [*Le livre*] [*racontait l’histoire*] [*d’un CHOU **bleu***] [*qui souffrait*] [*de sa différence*]). We chose two words that begin with a voiceless consonant so that fundamental frequency (f0) information

began at vowel onset. To avoid coarticulation effects due to contextualized-speech, each word was extracted from its carrier sentences, and the five speakers then heard their own productions of the two target words (*chou* and *chaud*) in their stressed and unstressed versions in isolation, and were instructed to reproduce the different versions. We used these repeated versions to construct our stimuli. Acoustic analyses were conducted to ensure that the repeated words were correctly stressed. Preboundary lengthening and f0 rises, the two main correlates of stress in French, were measured. Since we know that preboundary lengthening affects preferentially the rime of the syllable for primary stress (Astésano, 2001) and that our target syllables did not contain coda, the vowel portion of the word adequately reflects pre-boundary lengthening in our data. The analyses revealed that, compared to unstressed words, the vowel nuclei of the stressed words were longer (76 vs. 131 ms,  $t(18)=8.94$ ,  $p<.0001$ ) and associated with higher f0 values (208 vs. 394 Hz;  $t(18)=5.53$ ,  $p<.0001$ ). Also, an automatic detection of f0 maxima and minima showed that the vowel nuclei of the two target words *chaud* and *chou* were associated with an f0 rise when these targets bore primary stress (average percentage of rise between the minima and maxima: 64%) while the vowel nuclei were associated with an f0 plateau when the targets were unstressed (average percentage of rise between the minima and maxima: 0%). The vocalic durations, f0 maxima and minima and percentage of rise for each of the speaker are given in Table 1. As shown in that Table, the same word (e.g., *chaud*) in its stressed version was produced with various degrees of lengthening by the different speakers (e.g., vocalic duration for the male speaker = 151 ms; vocalic duration for the female speaker 1 = 142 ms), as well as with different degrees of pitch rises (e.g., pitch rise for the male speaker = 73%; pitch rise for the female speaker 1 = 63%).

Table 1. Acoustic properties of the two target words in their stressed and unstressed versions for the five speakers.

			Vocalic duration (ms)	f0 minima* (Hz)	f0 maxima (Hz)	Pitch rise (%)
Male speaker	<i>chaud</i>	stressed	151	140	242	73
	<i>chaud</i>	unstressed	94	99	103	4
	<i>chou</i>	stressed	142	125	247	98
	<i>chou</i>	unstressed	80	112	108	-4
Female speaker 1	<i>chaud</i>	stressed	142	295	480	63
	<i>chaud</i>	unstressed	67	238	242	2
	<i>chou</i>	stressed	140	341	488	43
	<i>chou</i>	unstressed	70	244	242	-1
Female speaker 2	<i>chaud</i>	stressed	123	241	393	63
	<i>chaud</i>	unstressed	87	224	220	-2
	<i>chou</i>	stressed	126	282	429	52
	<i>chou</i>	unstressed	73	218	211	-3
Female speaker 3	<i>chaud</i>	stressed	154	233	479	106
	<i>chaud</i>	unstressed	65	221	224	1
	<i>chou</i>	stressed	106	252	350	39
	<i>chou</i>	unstressed	76	226	229	1
Female speaker 4	<i>chaud</i>	stressed	110	279	448	61
	<i>chaud</i>	unstressed	67	239	239	0
	<i>chou</i>	stressed	116	273	382	40
	<i>chou</i>	unstressed	79	258	258	0

\* f0 minima and f0 maxima labels are only appropriate for stressed words, for unstressed

words it corresponds to the onset and offset of the vowel.

To match our stimuli on the initial /ʃ/ phoneme, we also asked the five speakers to produce the word *chat* ‘cat’ using the same procedure as above. For each speaker, we extracted the /ʃ/ from the word *chat* ‘cat’ in its unstressed version and lengthened it to a value that was the average of its duration in its stressed and unstressed version (110 ms for the male speaker). We then pasted onto this initial /ʃ/ the vowel /u/ and the vowel /o/ in both their stressed and unstressed versions to obtain a total of 4 stimuli for each of the 5 speakers which differed only in the acoustic characteristics of the vowel. To avoid clicks or audible discontinuities, the splicing and lengthening of the /ʃ/ were done at zero-crossings in parts of the /ʃ/ that were as

stable as possible, using PSOLA resynthesis routine (Pitch Synchronous Overlap and Add ; Moulines & Charpentier, 1990) on Praat (Boersma & Weenink, 2015). Figure 1 shows the words *chaud* ‘hot’ and *chou* ‘cabbage’ in their unstressed and stressed versions for the male speaker.

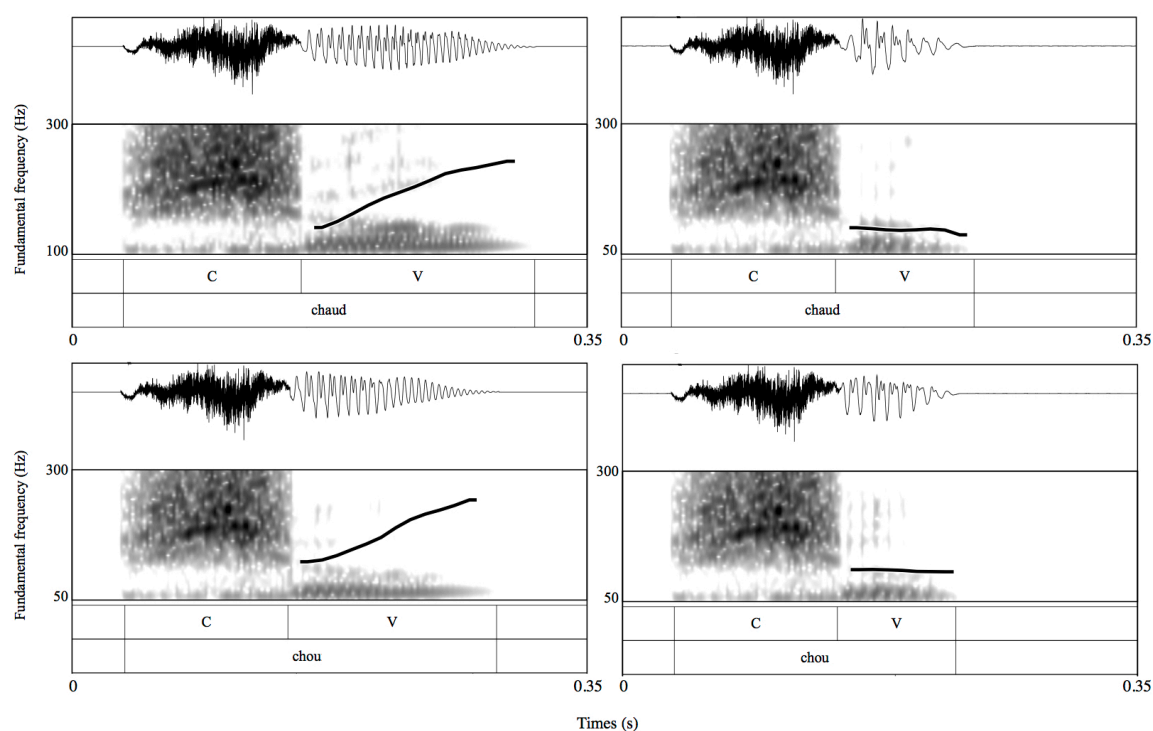


Figure 1. Phonemic and prosodic profile for the re-synthesized target words *chaud* ‘hot’ and *chou* ‘cabbage’ produced by the male voice in the stressed (left) and unstressed (right) contexts.

### C. Procedure

Participants sat in front of a computer screen in an acoustically shielded chamber. Stimuli were presented binaurally via Seinnheiser PC310 headphones, and the volume for auditory stimuli was adjusted to a comfortable level which was the same for all participants (60 dBA). The sound card of the computer used to display stimuli was a Creative SB-X-Fi. Experimental trials consisted of five monosyllabic words, separated by a 600 ms stimulus onset

asynchrony (SOA). On each trial, the same word was pronounced four times in either its stressed or unstressed form by four different female speakers. The order of the female voices was counterbalanced across the trials. The last stimulus, the target word, was produced by a male speaker and was either identical to (*identical* condition), or different from (*deviant* conditions) the first four words. Table 2 illustrates the experimental conditions. 96 trials were presented per condition. So that there were an equal number of trials requiring a “same” and “different” response, 192 filler trials were added in the identical condition but were not included in the analyses (i.e. 48 /**ju**'-/ /**ju**'/ trials, 48 /**ju**-/ **ju**/, 48 /**jo**'-/ /**jo**'/, 48 /**jo**-/ /**jo**/). The 192 filler trials were coded as “fillers” before the running of the experiment, and presented randomly among all other experimental trials. Each participant heard a total of 576 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 stress as the four preceding words, and to respond different when the target word differed from the preceding words either in its sounds, in its rhythm/melody or both in its sounds and rhythm/melody. The exact instructions are given in the APPENDIX. The buttons for making the “same” or “different” responses were counterbalanced across participants. Reaction times (RTs) were measured from the onset of the target word. An inter-trial interval of 3000 ms elapsed between the end of one trial and the beginning of the next. A training session comprised of 24 trials representative of the balance between conditions in the test trials was used to familiarize the participants with the task. During the experimental phase, participants received four blocks of 144 trials separated by a short break, and no feedback was given to participants. The experiment lasted one hour.

Table 2. Stimuli used in each condition.

<i>Conditions</i>	<i>Words produced by female speakers</i>	<i>Target word produce by the male speaker</i>	<i>Number of trials</i>	<i>Expected response</i>
<b><i>Pros&amp;phon identical</i></b>	/fʊ' /- /fʊ' /- /fʊ' /- /fʊ' /	/fʊ' /	24 trials (+48 fillers)	same
	/fʊ /- /fʊ /- /fʊ /- /fʊ /	/fʊ /	24 trials (+48 fillers)	same
	/fɔ' /- /fɔ' /- /fɔ' /- /fɔ' /	/fɔ' /	24 trials (+48 fillers)	same
	/fɔ /- /fɔ /- /fɔ /- /fɔ /	/fɔ /	24 trials (+48 fillers)	same
<b><i>Pros dev</i></b>	/fʊ /- /fʊ /- /fʊ /- /fʊ /	/fʊ' /	24 trials	different
	/fʊ' /- /fʊ' /- /fʊ' /- /fʊ' /	/fʊ /	24 trials	different
	/fɔ /- /fɔ /- /fɔ /- /fɔ /	/fɔ' /	24 trials	different
	/fɔ' /- /fɔ' /- /fɔ' /- /fɔ' /	/fɔ /	24 trials	different
<b><i>Phon dev</i></b>	/fɔ' /- /fɔ' /- /fɔ' /- /fɔ' /	/fʊ' /	24 trials	different
	/fɔ /- /fɔ /- /fɔ /- /fɔ /	/fʊ /	24 trials	different
	/fʊ' /- /fʊ' /- /fʊ' /- /fʊ' /	/fɔ' /	24 trials	different
	/fʊ /- /fʊ /- /fʊ /- /fʊ /	/fɔ /	24 trials	different
<b><i>Pros&amp;phon dev</i></b>	/fɔ /- /fɔ /- /fɔ /- /fɔ /	/fʊ' /	24 trials	different
	/fɔ' /- /fɔ' /- /fɔ' /- /fɔ' /	/fʊ /	24 trials	different
	/fʊ /- /fʊ /- /fʊ /- /fʊ /	/fɔ' /	24 trials	different
	/fʊ' /- /fʊ' /- /fʊ' /- /fʊ' /	/fɔ /	24 trials	different

#### D. Data acquisition and analysis

The electroencephalogram was recorded from the scalp with a 64-channel BioSemi Active-Two system. Individual electrodes were adjusted to a stable offset lower than 20 mV. The EEG signals were digitized at 512 Hz, re-referenced offline to the average of left and right mastoids and were bandpass filtered offline (0.4-30Hz). Independent Component Analysis decomposition was used to remove the artefacts generated by eye blinks and saccades (Makeig, Bell, Jung & Sejnowski, 1996). The epochs, starting 100 ms before target word onset (baseline) and ending 1000 ms after, were averaged for each condition and for each participant.

After visual inspection of the grand average of all conditions, four time windows were chosen for statistical analysis around the peak amplitude of the four components identified: 140-180 (N100), 180-220 (P200), 260-350 (N200), 700-900 (LPC). Since our target words were matched on their initial /ʃ/ phoneme thanks to the cross-splicing, we fixed the latencies for the EEG components relative to stimuli onset. Note that the corresponding time-windows relative to deviance onset were: 30-70 (N100), 70-110 (P200), 150-240 (N200), 590-790 (LPC). The scalp surface was divided into six regions of five electrodes each: frontal left (AF7, AF3, F5, F3, F1), frontal right (AF8, AF4, F6, F4, F2), central left (FC5, FC3, FC1, C3, C1), central right (FC6, FC4, FC2, C4, C2), occipito-parietal left (P5, P3, P1, PO7, PO3) and occipito-parietal right (P6, P4, P2, PO8, PO4).

Repeated measures analysis of variance (RM-ANOVA) was used for statistical assessment of both behavioral and ERP data, including condition (*pros&phon identical*, *pros dev*, *phon dev* and *pros&phon dev*) as the main factor. The analysis of ERP data additionally included site (frontal/central/occipito-parietal) and laterality (left/right) to model the topographical distribution of the effects. All p-values reported below were adjusted using the Greenhouse-Geisser correction for nonsphericity when appropriate, and a Bonferroni correction was used in post-hoc comparisons.

### **III. RESULTS**

#### **A. Behavioral results**

Correct discrimination rates and their corresponding d-prime (d') values are shown in Table 3. d' values from the accuracy data were calculated using the function  $(\text{hit}, \text{fa}) \{ \text{qnorm}(\text{hit}) - \text{qnorm}(\text{fa}) \}$  of the statistical software R (<http://www.r-project.org>) (see Pallier, 2002). Responses were coded as hits when participants correctly responded “different”, and as false alarms when participants incorrectly responded “different” on the “same” trials (i.e., 16 % of false alarms). An analysis of variance was run on the d' data with condition (*pros dev*, *phon dev* and *pros&phon dev*) as within factor. The main effect of condition was significant. Post-hoc t-tests with a Bonferroni correction revealed greater sensitivity in the *pros&phon dev* condition compared to the *pros dev* and the *phon dev* conditions. No significant difference was found between the *pros dev* condition and the *phon dev* condition. P-values (Greenhouse-Geisser corrected for the main effect of condition and Bonferroni corrected for post-hoc comparisons) are given in Table 4.

Table 3. Correct discrimination rates and their corresponding d-prime values, and mean RTs (in ms) for the three deviant conditions. Standard deviations are given in parenthesis.

	<i>Pros&amp;phon identical</i>	<i>Pros dev</i>	<i>Phon dev</i>	<i>Pros&amp;phon dev</i>
Hits	84	79	80	94
d' prime		2.07	2.20	2.93
RTs	1159 (210)	1167 (173)	1064 (191)	1017 (191)



Table 4. ANOVA table reporting p values (Greenhouse Geisser corrected for the main effect of condition and Bonferroni corrected for post-hoc contrasts) for ERP data.

	<b>d' values for deviant conditions</b>		<b>RTs for deviant conditions</b>
<b>Main effect of condition</b>	F(2,32) = 21.95 p<.0001 *	<b>Main effect of condition</b>	F(2,32) = 19.72 p < .0001 *
<b>Bonferroni Post-Hoc</b>		<b>Bonferroni Post-Hoc</b>	
Pros dev = phon dev	p = 1.00	Pros dev > phon dev	p<.001 *
Pros dev < pros&phon dev	p<.0001 *	Pros dev > pros&phon dev	p<.0001 *
Phon dev < pros&phon dev	p<.0001 *	Phon dev = pros&phon dev	p=.19

Because participants undertake very many trials, an additional analysis on d-prime values was conducted to examine how the effects especially in the *pros dev* condition developed over the course of the experiment. D-prime values in each condition were divided in four blocs as a function of their position in the experimental list, and we compared the performance between the first and the last bloc of the experiment. Neither a main effect of bloc ( $F(1,16)=0.29$ ,  $p>.20$ ), nor a significant interaction between bloc and condition was found ( $F(2,32)=1.51$ ,  $p>.20$ ), in contrast to what would have been expected if participants had learned to process stress information over the course of the experiment.

RTs for “different responses” in deviant conditions were entered into an analysis of variance. RTs longer than 2300 ms were removed from analyses. This criterion was based on the overall distribution of RTs: mean = 1193 ms, median = 987, SD = 693; it was set to 2300 which corresponds to the rounding inferior of median + 2 x SD (= 2373). Using this criterion 6% of the data were discarded. We found a main effect of condition. Post-hoc t-tests with a Bonferroni correction revealed that participants responded slower in the *pros dev* condition than in the *phon dev* condition. They also responded slower in the *pros dev* condition than in the

*pros&phon dev* condition. By contrast the difference between the *pros&phon dev* condition and the *phon dev* condition was not significant (see Table 4).

## **B. EEG Results**

In the N100 time window (140-180 ms), we only observed a main effect of sites [ $F(2,32)=4.40$ ,  $p < .05$ ], with central and occipito-parietal regions showing more negative values than frontal ones.

In the P200 time window (180-220 ms), a trend effect of condition was observed [ $F(3,48) = 2.82$ ,  $p = .05$ ] with more positive values in the *pros&phon identical* condition than in the *pros dev*, the *phon dev* and the *pros&phon dev* conditions. A significant laterality x condition interaction was found [ $F(3,48)=3.05$ ,  $p < .05$ ], due to a greater effect of condition over the left hemisphere compared to the right hemisphere. No significant difference was found between the three deviant conditions. Corrected p-values for main effects are given in Table 5. Post-hoc contrasts are given in Table 6. Figure 2 displays average waveforms in each condition in the left hemisphere in the three regions for which the maximum P200 responses were found.

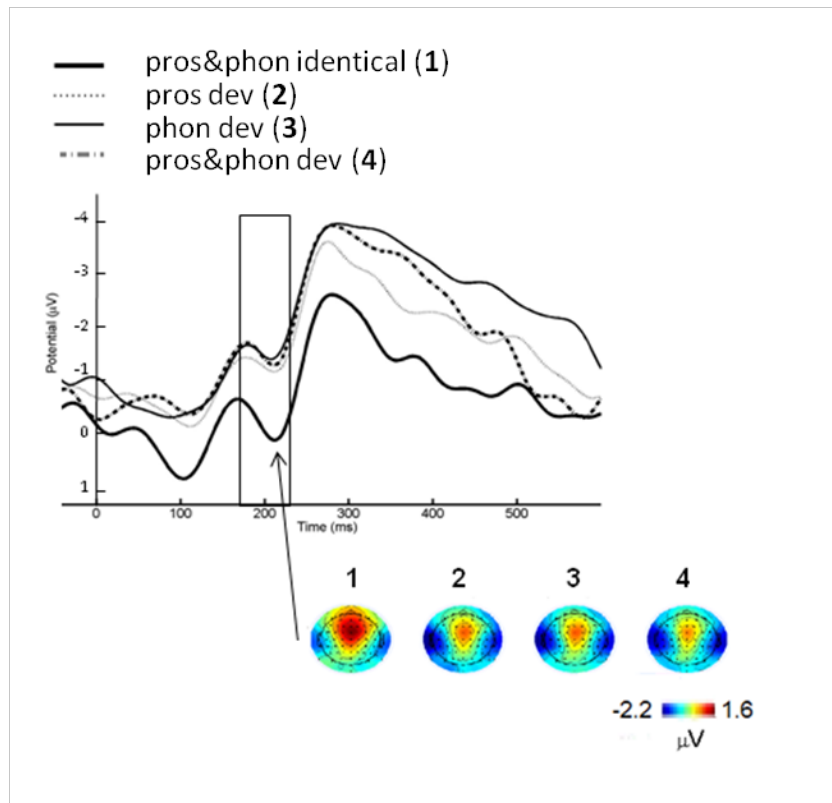


Figure 2. Averages ERPs at left hemisphere in the four conditions (*pros&phon identical, pros dev, phon dev, pros&phon dev*). The maps show the distribution of the P200 component in the 180-220 ms range window.

Table 5. ANOVA table reporting p values (Greenhouse Geisser corrected for all main effects) for ERP data.

	<b>P2</b>	<b>N2</b>	<b>LPC</b>
Main effect of condition	p=.05	p<.05*	p<.05*
Main effect of laterality	p=.24	p<.01*	p<.0001*
Main effect of site	p=.05	p<.01*	p<.0001*
Condition x laterality	p<.05*	p=.06	p=.10
Condition x site	p=.22	p<.05*	p<.0001*
Condition x laterality x site	p=.46	p=.31	p<.0001*

Table 6. Post-hoc contrasts (with Bonferroni correction). Sites at which components reach their maximal values are in bold.

	P2		N2			LPC					
	Left hemisphere	Right hemisphere	Frontal sites	Central sites	Occipito-parietal sites	Frontal left site	Frontal right site	Central left site	Central right site	Occipito-parietal left site	Occipito-parietal right site
Pros&phon identical vs. pros dev	<b>p&lt;.0001*</b>	p<.01 *	<b>p&lt;.0001*</b>	p<.001*	p<.05	p=1.00	p=1.00	p=1.00	p=.15	p=1.00	<b>p&lt;.0001*</b>
Pros&phon identical vs. phon dev	<b>p&lt;.0001*</b>	p=.08	<b>p&lt;.0001*</b>	p<.0001*	p<.0001*	p=.71	p=1.00	p=1.00	p=1.00	p=1.00	<b>p=1.00</b>
Pros&phon identical vs. pros&phon dev	<b>p&lt;.0001*</b>	p<.001 *	<b>p&lt;.0001*</b>	p<.0001*	p<.0001*	p=1.00	p=.10	p<.0001*	p<.001*	p=1.00	<b>p&lt;.0001*</b>
Pros dev vs. Phon dev	<b>p = 1.00</b>	p=1.00	<b>p&lt;.0001*</b>	p<.01*	p<.0001*	p=1.00	p=1.00	p=1.00	p=1.00	p=1.00	<b>p=.45</b>
Pros dev vs. Pros&phon dev	<b>p = 1.00</b>	p=1.00	<b>p&lt;.0001*</b>	p=.18	p<.05*	p=1.00	p=1.00	p<.0001*	p=1.00	p=1.00	<b>p=1.00</b>
Phon dev vs. Pros&phon dev	<b>p = 1.00</b>	p=1.00	<b>p=1.00</b>	p=1.00	p=1.00	p=1.00	p<.01*	p<.0001*	p<.05*	p<.05*	<b>p=.21</b>

In the N200 time window (260-350 ms), there was a significant main effect of condition [ $F(3,48) = 4.13, p < .05$ ]. The *pros&phon dev* and *phon dev* conditions had more negative values than the *pros dev* condition, which had in turn more negative values than the *pros&phon identical* condition. No significant difference was found between the *pros&phon dev* condition and the *phon dev* condition. A significant interaction between the condition and the site was observed [ $F(6,96) = 2.52, p < .05$ ], due to greater effect of condition over frontal sites than central and occipito-parietal sites. Also, a trend interaction between condition and laterality [ $F(3,48) = 2.80, p = .06$ ] was found, due to greater effect of condition over left regions compared

to right regions. Corrected p-values for main effects are given in Table 5. Post-hoc contrasts are given in Table 6. Figure 3 illustrates average waveforms in all conditions at frontal sites over which the N200 component reaches its maximal values.

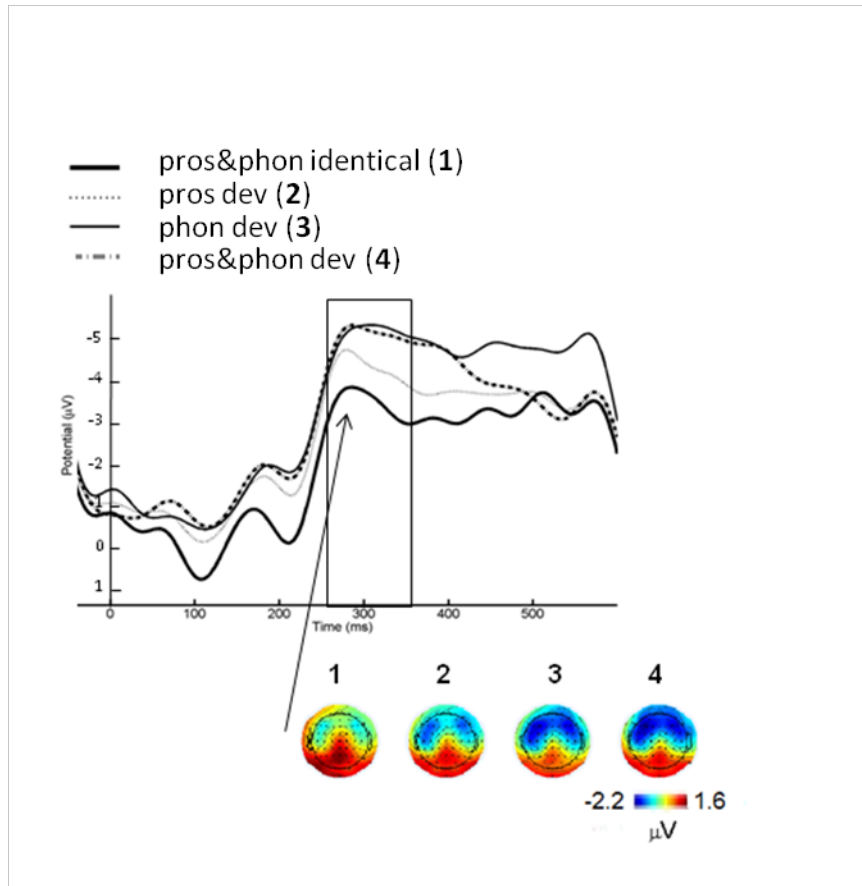


Figure 3. Averages ERPs at frontal sites in the four conditions (*pros&phon identical*, *pros dev*, *phon dev*, *pros&phon dev*). The maps show the distribution of the N200 component in the 260-350 ms range window for each condition.

In the late time window (700-900 ms), there was a significant main effect of condition [ $F(3,48) = 3.23, p < .05$ ]. Larger positivity for the *pros&phon dev* condition compared to the *phon dev* condition was observed. A significant condition x site interaction was found [ $F(6,96) = 14.91, p < .0001$ ], due to greater effect of condition over occipito-parietal sites than central and frontal sites. Post-hoc comparisons testing this interaction showed larger positivity for the

*pros dev* and the *pros&phon dev* condition compared to the *pros&phon* identical and the *phon dev* conditions. A significant condition x laterality x site interaction was also found, due to a greater effect of condition at occipito-parietal right sites compared to other sites ( $[F(6,96) = 10.8, p < .0001]$ ). At occipito-parietal right sites more positive values were found for the *pros dev* and *pros&phon dev* conditions compared to the *pros&phon* identical condition whereas the difference between the *pros dev* and *pros&phon dev* conditions compared to the *phon dev* condition was not significant. Corrected p-values for main effects are given in Table 5. Post-hoc contrasts are given in Table 6. Figure 4 illustrates averages waveforms for all conditions at occipito-parietal right sites

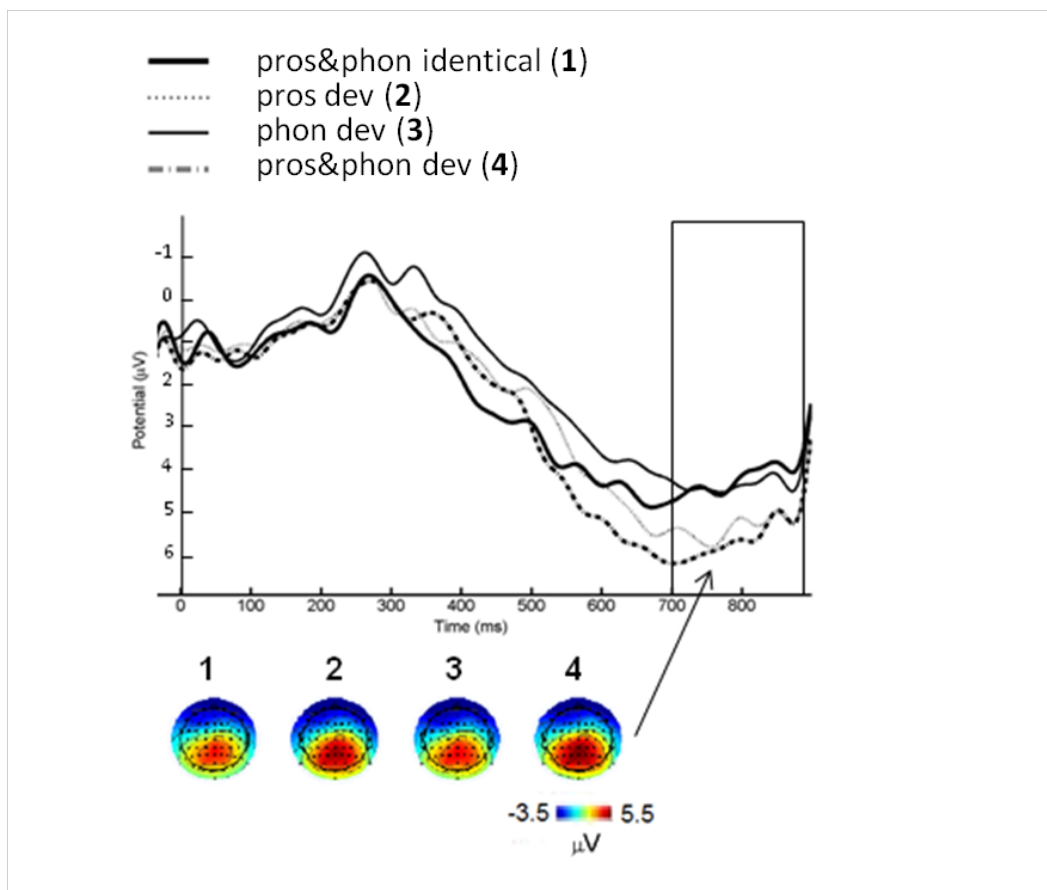


Figure 4. Averages ERPs at occipito-parietal right regions in the four conditions (*pros&phon* identical, *pros dev*, *phon dev*, *pros&phon dev*). The maps show the distribution of the LPC component in the 700-900 ms range window.

## DISCUSSION

The main goal of this experiment was to examine whether French participants can use stress information at a phonological level when asked to discriminate between words that differ in the presence vs. absence of primary stress. To prevent French listeners from relying on an acoustically-based response strategy and to encourage them to extract an abstract representation of stress, we used acoustically variable stimuli that were produced by different speakers. Crucially, we observed significant differences on both the P200 and N200 components when the target word (e.g., /*ju*/) differed from the preceding word only in stress (e.g., /*ju'*/) in comparison to a control condition in which the target word and the preceding word had the same stress pattern and the same phonemes (e.g., /*ju*/). This finding provides strong evidence that French participants are sensitive to stress differences, at least when they have to discriminate between stressed and unstressed words.

Because we used monosyllabic words that were either stressed or unstressed, it remains possible that our participants have distinguished the target word from the preceding word by comparing the relative acoustic differences between stressed and unstressed syllables and this despite the fact that we used different speakers. However, since the first response related to the detection of a prosodic mismatch was seen on the P200 component which is generally interpreted as reflecting phonological processes (Dehaene-Lambertz, 1997; Dehaene-Lambertz et al., 2000; Brunellière et al., 2009), we believe that our participants have processed stress differences at a higher level of processing than the acoustic level. This claim is also supported by the fact that we did not observe an effect in the case of a stress mismatch on the N100 component classically associated with acoustical processing. Interestingly, the analyses performed on the P200 component also revealed no significant difference between the stress

deviant (/ʒu' /- /ʒu/) and the phoneme deviant (/ʒo /- /ʒu/) conditions. This suggests that early in the identification of words, acoustic information which is relevant to specifying the presence vs. absence of stress is transformed into a more abstract representation, much like phoneme-relevant acoustic information is transformed into abstract phoneme categories. Hence, it appears that French listeners can rely on abstract properties of stress to categorize a word as stressed or unstressed, like they are able to rely on abstract properties of a phoneme to categorize a vowel as being /o/ or /u/.

In contrast to what we observed on the P200 time window (180-220 ms after target onset), the analyses performed on a later time window corresponding to the N200 component (260-350 ms after target onset) showed a significant difference between stress and phonemic contrasts. In particular, a response with greater N200 amplitudes was observed for a mismatch on the last phoneme (/ʒo /- /ʒu/) in comparison to a mismatch in the stress pattern (/ʒu' /- /ʒu/). This suggests that as French listeners progress in their processing of words, they react more strongly to phonemic than to stress information. As a consequence, changes in phonemic information lead to a stronger electrophysiological response than changes in stress information. This stronger response in case of phonemic mismatch as compared to a prosodic mismatch was reflected in our RTs data since the participants were faster in the phoneme deviant than in the stress deviant condition. Together our N200 and RTs results corroborate Dupoux et al. (1997) in showing processing difference between stress and phonemic contrast, even if this difference was not reflected in our error rates.

Finally, we were also interested in examining how stress information combines with phonemic information. To this end, we tested a condition in which the target word (/ʒu/) differed from the preceding word on both stress and phoneme make-up (/ʒo' /). On both the P200 and



N200 time windows, no difference was found between this latter condition and the condition in which the target word (/ʃu/) differed from the preceding word on only the last phoneme (/ʃo/). These ERP findings are in line with RTs data that also showed no significant difference between these two conditions. Thus as in Dupoux et al. (1997, 2001), French participants did not benefit from the presence of stress information in discriminating words when these words already differed in phonemic information.

In accordance with others studies using the same paradigm, we reported a LPC component, in a late time window, which is generally elicited when participants have to discriminate between two words, and which has been interpreted as reflecting the conscious detection of a deviant stimuli compared to the preceding context stimuli (Dehaene et al., 1997; Dehaene, 2000; Brunellière et al., 2009). Analyses on this late component revealed greater amplitudes when the target word (/ʃu/) differed from the preceding word on stress (/ʃu'/) and when it differed both on stress and on the last phoneme (/ʃo'/), in comparison to the control condition in which the target word (/ʃu/) share both stress and phonemic information with the preceding word (/ʃu/). Hence, significant differences emerged in this late time window when target words differed from the preceding word on the stress. The LPC has been shown to be influenced by attentional factors (Li, Zinbarg, Boehm, Paller, 2008). It is possible that in order to perform the task, our participants paid more attention to prosodic differences which could explain why significant differences emerged on the LPC component in case of prosodic mismatch.

To sum-up, our study shows that French listeners are sensitive to stress differences when they have to distinguish between two words that differ in the presence vs. absence of primary stress (/ʃu/ vs. /ʃu'/) in their own language. It appears thus that French listeners are capable of

creating a representation of stress that is more abstract than previously observed in behavioral studies (Dupoux et al., 1997). A crucial difference between our study and that of Dupoux et al. is that we used monosyllabic items, whereas Dupoux et al. used tri-syllabic items with a large variety of consonants and vowels. Hence despite the fact that we introduced phonetic variability by mixing different speakers in our stimuli, our task could be easier than that in Dupoux et al.'s study. Nonetheless, despite this difference in the length of the items, our study suggests that in contrast to the claim made by Dupoux et al. that French listeners are not deaf to stress. Note that such a conclusion has been recently supported by Schwab & Llisteri (2014)'s study showing that it is possible to train French listeners to perceive stress contrasts in tri-syllabic Spanish items.

To conclude, the present results converge in showing that French listeners can rely on abstract properties of stress to categorize a word as stressed or unstressed. Pre-lexical processing may thus involve abstraction processes not only for segmental/phonemic information (McQueen, Cutler & Norris, 2006) but also for suprasegmental features such as stress (Sulpizio & McQueen, 2001) even in French, a language in which stress is not lexically contrastive. The precise mechanisms underlying this abstraction process are likely to be dependent on the native language of the listeners. For example, we expect that in the same kind of paradigm, Spanish or English listeners for whom stress is lexically contrastive would not show any differences between the stress and phoneme contrasts on the N200 component. Additional cross-linguistic studies are thus required to examine in more detail this prelexical abstraction process, and how it varies as a function of the prosodic characteristics of languages.

## **ACKNOWLEDGMENTS**

This work was supported by the Labex Brain and Language Research Institute (ANR-11-LABX-0036) and has benefited from support from the French National Agency of Research (ANR), under the project title 'Investments of the Future' A\*MIDEX (ANR-11-IDEX-0001-02). This research was also supported by Grant 100014\_131954 from the Swiss National Science Foundation to Ulrich Frauenfelder.

## REFERENCES

- Astésano, C. (2001). *Rythme et accentuation en français: invariance et variabilité stylistique* (l'Harmattan, Paris), 337 pp.
- Banel, M. H., and Bacri, N. (1994). "On metrical patterns and lexical parsing in French", *Speech Com.* **15**(1), 115-126.
- Boersma, P., and Weenink, D. (2015). "Praat. doing phonetics by computer" (Version 5.4.01)(2015), Computer program: [www.praat.org](http://www.praat.org) (date last viewed 30/04/15).
- Brunellière, A., Dufour, S., Nguyen, N., and Frauenfelder, U. H. (2009). "Behavioral and electrophysiological evidence for the impact of regional variation on phoneme perception", *Cognition* **111**(3), 390-396.
- Christophe, A., Peperkamp, S., Pallier, C., Block, E., and Mehler, J. (2004). "Phonological phrase boundaries constrain lexical access I. Adult data", *J. Memory and Lang.* **51**(4), 523-547.
- Dehaene-Lambertz, G. (1997). "Electrophysiological correlates of categorical phoneme perception in adults", *NeuroReport* **8**(4), 919-924.
- Dehaene-Lambertz, G., Dupoux, E., and Gout, A. (2000). Electrophysiological correlates of phonological processing: a cross-linguistic study, *J. Cognitive Neuro.* **12**(4), 635-647.
- Dupoux, E., Pallier, C., Sebastian, N., and Mehler, J. (1997). "A distressing "deafness" in French?", *J. Memory and Lang.* **36**(3), 406-421.
- Dupoux, E., Peperkamp, S., and Sebastián-Gallés, N. (2001). "A robust method to study stress "deafness"", *J. Acoust. Soc.* **110**(3), 1606-1618.
- Dupoux, E., Peperkamp, S., and Sebastián-Gallés, N. (2010). "Limits on bilingualism revisited: Stress 'deafness' in simultaneous French-Spanish bilinguals", *Cognition* **114**(2), 266-275.
- Eulitz, C., & Lahiri, A. (2004). Neurobiological evidence for abstract phonological representations in the mental lexicon during speech recognition. *Journal of Cognitive*

*Neuroscience*, **16**, 577–583.

- Li, W., Zinbarg, R. E., Boehm, S. G., & Paller, K. A. (2008). Neural and behavioral evidence for affective priming from unconsciously perceived emotional facial expressions and the influence of trait anxiety. *Journal of Cognitive Neuroscience*, **20**, 1-13.
- Makeig, S., Bell, A. J., Jung, T.-P., & Sejnowski, T. (1996). Independent component analysis of electroencephalographic data. In: D. S. Touretzky, M. C. Mozer, & M.E. Hasselmo (Eds.), *Advances in neural information processing systems* (Cambridge: MA: MIT Press, pp. 145-151.
- McQueen, J. M., Cutler, A., & Norris, D. (2006). Phonological abstraction in the mental lexicon. *Cognitive Science* **30**(6), 1113-1126.
- Moulines, E. and Charpentier, F. (1990). “Pitch-synchronous wave- form processing techniques for text-to-speech synthesis using diphones”, *Speech Com.* **9**(5), 453–467.
- Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical Neurophysiology* **118**(12), 2544-2590.
- Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R. J., Luuk, A., Allik, J., Sinkkonen, J. and Alho K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, **385**, 432–434.
- Näätänen, R., Simpson, M. and Loveless, N. E. (1982). Stimulus deviance and evoked potentials. *Biological Psychology* **14**, 53–98.
- Pallier, C. (2002). Computing discriminability and bias with the R software. URL <http://www.pallier.org/ressources/aprime/aprime>.

- Peperkamp, S., Vendelin, I., and Dupoux, E. (2010). "Perception of predictable stress: A cross-linguistic investigation", *J. Phon.* **38**(3), 422-430.
- Rietveld, A. C. M. (1980). "Word boundaries in the French language", *Language and Speech* **23**, 289–296.
- Schwab, S., & Llisterri, J. (2014). Does training make French speakers more able to identify lexical stress? *COPAL, Concordia Working Papers in Applied Linguistics*, 5, 624–636.
- Spinelli, E., Welby, P., and Schaegis, A. L. (2007). "Fine-grained access to targets and competitors in phonemically identical spoken sequences: the case of French elision", *Language cogn. proc.* **22**(6), 828-859.
- Spinelli, E., Grimault, N., Meunier, F., and Welby, P. (2010). An intonational cue to word segmentation in phonemically identical sequences, *Attention, Perception, & Psychophysics* **72**(3), 775-787.
- Sulpizio, S., & McQueen, J. M. (2012). Italians use abstract knowledge about lexical stress during spoken-word recognition. *Journal of Memory and Language* **66**, 177-193.

## APPENDIX

### *Instructions given to participants (in French and translated):*

#### **Votre tâche :**

Au cours de cette expérience, vous allez entendre quatre mots prononcés par des voix de femmes. Il s'agira tout le temps du même mot prononcé avec les mêmes sons et la même mélodie/rythme.

Le dernier mot que vous allez entendre sera prononcé par une voix d'homme. C'est sur ce mot que vous devrez réaliser une tâche.

Vous devez appuyer sur la touche « MÊME » si le mot prononcé par la voix d'homme comporte les mêmes sons prononcés avec la même mélodie/rythme que les mots prononcés par les voix de femmes.

Vous devez appuyer sur « DIFFERENT » si le mot prononcé par la voix d'homme ne comporte pas les mêmes sons et/ou la même mélodie/rythme que les mots prononcés par les voix de femmes.

Nous allons écouter trois exemples :

Dans le premier exemple, le mot « chou » prononcé par la voix d'homme est strictement identique au mot prononcé par les voix de femmes (mêmes sons et même mélodie/rythme).  
ECOUTER l'EXEMPLE 1

Dans le deuxième exemple, le mot prononcé la voix d'homme (« chou ») ne comporte pas les mêmes sons que le mot « chaud » prononcés par les voix de femme. Il est donc différent.  
ECOUTER l'EXEMPLE 2

Dans le troisième exemple, le mot « chou » prononcé par la voix d'homme n'est pas prononcé avec la même mélodie/rythme que le mot « chou » prononcé par les voix de femme. Il est donc différent.  
ECOUTER l'EXEMPLE 3

#### ***Your task:***

*During this experiment, you will hear four words spoken by women's voices. These words will always correspond to the same word pronounced with the same sounds and the same melody / rhythm.*

*The last word you will hear will be produced by a man's voice. You will have to perform a task based on this word.*

*You should press the "SAME" button if the word spoken by the man's voice has the same sounds pronounced with the same melody / rhythm as the words spoken by the women's voices.*

*You should press the "DIFFERENT" button if the word spoken by the man's voice does not have the same sounds and / or the same melody / rhythm as the words spoken by the women's voices.*

*We will listen to three examples:*

*In the first example, the word "chou" /ʃu/ spoken by the man's voice is identical to the word spoken by the women's voices (same sounds and same melody / rhythm).*

*LISTEN TO EXAMPLE 1*

*In the second example, the word spoken by the man's voice "chou" /ʃu/ does not have the same sounds as the word spoken by the woman's voice "chaud" /ʃo/. It is therefore different.*

*LISTEN TO EXAMPLE 2*

*In the third example, the word "chou /ʃu/" spoken by the man's voice is not pronounced with the same melody / rhythm as the word "chou /ʃu/" pronounced by the woman's voice. It is therefore different.*

*LISTEN TO EXAMPLE 3*