

Running Head: Accentual contrast discrimination

When native contrasts are perceived as non-native: the role of the ear of presentation in the  
discrimination of accentual contrasts

Amandine Michelas and Sophie Dufour

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

Amandine Michelas  
Laboratoire Parole et Langage  
Aix-Marseille Université, CNRS  
5, avenue Pasteur  
13604 Aix-en-Provence  
France  
Email: michelas@lpl-aix.fr

Abstract: In two ABX experiments using natural and synthetic stimuli, we examined the ability of French listeners to perceive accentual variation by manipulating the ear of presentation. A native (/balɔ̃/-/ba'ɪɔ̃/) and a non-native (/balɔ̃/-/ba'ɪɔ̃/) contrasts were tested. The stimuli A and B varied either in accent (/balɔ̃/-/ba'ɪɔ̃/), in one phoneme (/ba'lo/-/ba'ɪɔ̃/ “bundle-ball”) or in both types of information (/balɔ̃/-/ba'ɪɔ̃/). For the non-native contrast, persistent difficulty was found regardless of the ear of presentation. For the native contrast, better performance was observed when the stimuli were presented to the left ear, and thus when the processing was pushed into the right hemisphere. Our findings also showed that the native contrast was processed as a non-native contrast when the processing was pushed into the left hemisphere. More generally, our study indicates that accentual variation at the word level in French is processed as non-linguistic variation.

Keywords: Speech perception, right hemisphere advantage, ear of presentation, accentual variation, French prosody.

In lexical stress languages and in tone languages, prosodic information is linguistically contrastive at the word level. As a result, in these languages, prosodic information can change the meaning of a word as in the case of *BEbe* /'bebe/ “s/he drinks” and *beBE* /be'be/ “baby” in Spanish. Consequently, speakers of these languages show no difficulty in discriminating two items differing in prosodic information (e.g., Dupoux, Pallier, Sebastián & Mehler, 1997; Dupoux, Peperkamp & Sebastián-Gallés, 2001; Dupoux, Sebastián-Gallés, Navarrete & Peperkamp, 2008). Moreover, neuropsychological studies conducted on tone languages have shown that the left hemisphere is recruited to process contrastive prosodic information at the word level for speakers of these languages (e.g., Gandour, Wong, & Hutchins, 1998; Gandour et al., 2000, 2002, 2004; Klein, Zatorre, Milner, & Zhao, 2001). But what happens for languages such as French in which prosodic information does not create meaning distinctions at the word level?

One of the most striking characteristics of French prosody is that accent<sup>1</sup>, the acoustic prominence given to a syllable, is determined post-lexically. As a result, a same word receives accent or not depending on its location within the 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. Hence the word *ballon* conserves the same meaning whether it is accented on its final syllable or not<sup>2</sup>. Because accent is not lexically contrastive in French, it has been repeatedly shown that French listeners have difficulty discriminating between two items differing in lexical stress (e.g., *BEbe* /'bebe/ and *beBE* /be'be/; Dupoux et al., 1997; Dupoux et al. 2001; Dupoux et al. 2008). The term “stress deafness” has been widely used to qualify the low sensitivity of French listeners to accentual variation affecting the word level. However, recent studies conducted in our laboratory (Michelas, Frauenfelder, Schön & Dufour, 2016; Michelas,

Esteve-Gibert & Dufour, 2018) have shown that this low sensitivity to accentual variation does not extend to all accentual contrasts. In fact, we showed that French listeners are fully able to discriminate between an unaccented (e.g., *ballon* /balɔ̃/ “ball”) and an accented word (*baLLON* /ba'ɫɔ̃/ “ball”), namely between accentual differences that exist in their native language. Based on past studies showing that other sources of speech variation (i.e., talker variation) are processed by the right hemisphere when they are not linguistically relevant (Van Lancker & Canter, 1982; Von Kriegstein, Eger, Kleinschmidt & Giraud, 2003; Gonzalez & McLennan, 2007), we hypothesized that accentual variation at the word level in French would be primarily processed by the right hemisphere. In particular, using dichotic presentation, we examined the role of the ear of presentation and we predicted better performance in the discrimination of both native and non-native accentual contrasts when words are presented to the left ear, and thus, due to contralateral projections, when the processing was pushed into the right hemisphere.

What do we already know about French listeners' ability to discriminate accentual contrasts? In a seminal study, Dupoux and collaborators (Dupoux et al., 1997) compared French and Spanish listeners' ability to discriminate lexical stress contrasts that exist in Spanish but not in French. 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 to B, they showed that, compared to Spanish speakers, French speakers had more difficulty discriminating between two nonsense words that differed only in stress (FI<sub>d</sub>ape, fiDA<sub>p</sub>e). In addition, unlike Spanish listeners, French listeners performed worse when the stimuli differed in stress (FI<sub>d</sub>ape, fiDA<sub>p</sub>e) than in phonemic identity (FI<sub>d</sub>ape, LI<sub>d</sub>ape). Together, these findings suggest that French listeners process an accentual and a phonemic contrast differently. However, their difficulties could be due to the fact that Dupoux et al. (1997) used stress location contrasts (BO<sub>p</sub>elo/boPE<sub>l</sub>o and boPE<sub>l</sub>o/bopeLO) that do not exist in French.

Contrary to Dupoux et al. (1997), recent studies (Michelas et al., 2016; Michelas et al., 2018) have examined the ability of French listeners to discriminate accentual contrasts that are frequently encountered in their native language. Using the same paradigm, the authors replicated Dupoux et al. (1997) 's results showing that French listeners had more difficulty discriminating two French words that differed in accent location (*JUry-juRY* “jury-jury”) compared to words that differed in one phoneme (*juRON-juRY* “swearword-jury”). However, and crucially, no difficulty was observed when the stimuli differed in the presence or absence of accent (*jury-juRY*), namely on the accentual contrast that exists their native language. Such observation suggests that French listeners have difficulty discriminating non-native accentual contrasts, but that their difficulty with accent disappear as soon as they have to discriminate accentual contrasts that do exist in their native language.

Because French listeners have been shown to be able to differentiate an accented word from its unaccented version (Michelas et al., 2016; Michelas et al., 2018), in a recent study (Michelas & Dufour, 2019), we have examined whether this accentual variation could be stored in the mental lexicon. In that study, participants heard a first block of stimuli (the priming block) followed by a second block (the target block) in which some of the words from the first block were repeated either with the same accentuation (*banDEAU-banDEAU* “headband-headband”) or with a different accentuation (*bandeau-banDEAU* “headband-headband”). In comparison to a control condition in which primes and targets were unrelated (*marron-banDEAU* “chestnut-headband”), a repetition priming effect of similar magnitude was observed in the same and different accentuation conditions. Such a result suggests that accentual variation is not stored in the mental lexicon even though French native listeners are fully able to distinguish an accented and an unaccented word. This observation mirrors results for segmental contrasts in a

second language showing that good performance in discrimination or categorization tasks does not necessarily extend to tasks involving lexical processing (Darcy, Daidone & Kojima, 2013; Díaz, Mitterer, Broersma & Sebastián-Gallés, 2012).

All the studies conducted on the French language so far could therefore suggest that accentual variation at the word level is processed as non-linguistic variation by French listeners. In the current study, we examined this hypothesis following the approach of Gonzalez & McLennan (2007)'s study which showed that non-linguistic variation is primarily processed by the right hemisphere<sup>3</sup>. In particular, these authors showed that variation due to talker identity influenced spoken word processing only when stimuli were presented to the left ear, and thus primarily processed by the right hemisphere. Consequently, we wondered here whether French native listeners better succeed in discriminating accentual contrasts when stimuli were presented to the left ear. To do so, we used the Dupoux et al. (1997)'s ABX task again in which participants heard three words produced by three different speakers and had to decide whether X was similar to A or to B. A and B were produced by two different female speakers and X was always produced by a male speaker. Two types of accentual contrasts were tested: 1) the accentual contrast that exists in French, namely the difference between an unaccented word and the same word bearing primary accent on its last syllable: /balɔ̃/ and /ba'ɔ̃/ *ballon* "ball" (the 'native accentual pattern'), and 2) an accentual contrast that does not exist in French, namely the difference between a word bearing primary accent on its first syllable and a word bearing primary accent on its last syllable: /'balɔ̃/ and /ba'ɔ̃/, (the 'non-native accentual pattern'). Note that whatever the accentual contrast tested, accentual variation never creates meaning distinctions (i.e., *ballon* always means 'ball' whether it receives accent on its first syllable, on its second syllable or no accent) and thus the crucial manipulation concerns whether the accentual contrast exists or not in the French language. We expected participants to process an

accentual mismatch with the same ability as a phonemic mismatch when stimuli were presented to the left ear and thus primarily processed in the right hemisphere. Within each accentual contrast, the accentual mismatch was always compared to a phonemic mismatch. In accordance with the growing body of evidence showing that phonemic representations are bilateral (e.g., Hickok & Poeppel, 2000), phonemic mismatch is expected to be processed with the same level of performance whatever the ear of presentation. In addition, as in Dupoux et al.'s study, a redundant condition was also used in which stimuli differed both in the accentual and phonemic information. The experimental conditions are illustrated in Table 1.

Table 1. Experimental conditions.

	<b>Accentual mismatch</b>	<b>Phonemic mismatch</b>	<b>Accentual and phonemic mismatch</b>
<b>Native accentual pattern</b>	/balɔ̃/-/ba'ɪɔ̃/ <i>ballon</i> - <i>ballon</i> "ball" - "ball"	/ba'lo/-/ba'ɪɔ̃/ ballot - ballon "bundle" - "ball"	/balo/-/ba'ɪɔ̃/ ballot - ballon "bundle" - "ball"
<b>Non-native accentual pattern</b>	/'bɑlɔ̃/-/ba'ɪɔ̃/ <i>ballon</i> - <i>ballon</i> "ball" - "ball"	/'bɑlo/-/'bɑlɔ̃/ ballot - ballon "bundle" - "ball"	/'bɑlo/-/ba'ɪɔ̃/ ballot - ballon "bundle" - "ball"

## Experiment 1

In this experiment, native and non-native stimuli were obtained using natural speech.

### Method

Participants: Fifty right-handed French native speakers between 18 and 37 years old participated in the experiment. All were students from Aix-Marseille University and reported

having no neurological, hearing or language impairment. Each participant gave informed consent prior to experiment and was remunerated for their time. Half the participants heard stimuli in the right ear while the other half heard stimuli in the opposite ear.

Materials: Four pairs of 4-phoneme-long bisyllabic French words differing in one phoneme (e.g. *ballon* /balɔ̃/ “ball” vs. *ballot* /balo/ “bundle”) were selected. Three French speakers (two women, one man) produced the 8 target words (in bold in the following examples) within carrier sentences. Within these carrier sentences, target words either bore accent on their last syllable or were unaccented depending on their place within the phrase (e.g. [*On m’avait parLÉ*] [*d’un **ballon** maGIQUE*] [*qui ne se creVAIT*] [*presque jaMAIS*] “I was told about a magic ball that hardly ever burst” vs. [*On m’avait parLÉ*] [*d’un petit **baLLON***] [*qui ne se creVAIT*] [*presque jaMAIS*] “I was told about a little ball that hardly ever burst”). The same carrier sentences as for the unaccented condition were used to obtain 8 target words that bore accent on their first syllable (e.g. [*On m’avait parLÉ*] [*d’un **Ballon** maGIQUE*] [*qui ne se creVAIT*] [*presque jaMAIS*]). Because words bearing primary accent on the first syllable are not found in French, speakers were asked to produce accent on the first syllable of the target word by making the target syllable prominent. To avoid coarticulation effects due to contextualized speech, each word was extracted from its carrier sentence. The three speakers heard their own production of the 8 target words in their accented (both first accented syllable and second accented syllable) and unaccented versions in isolation, and were instructed to reproduce the different versions. The sentences and the 8 target words presented in isolation were recorded at a sampling rate of 44 100 Hz. Acoustic analyses using Praat software (Boersma & Weenink, 2020) were then conducted to ensure that the repeated words were produced with the expected accentual patterns. To do so, for the three versions of each word (unaccented, with a primary accent on the second syllable and with a primary accent on the first syllable), we measured the length of



both syllables as well as the f0 movement associated with these two syllables. The acoustic properties of the stimuli are shown in Table 2.

Table 2. Acoustic properties of the words in their unaccented and accented versions (1<sup>st</sup> and 2<sup>nd</sup> syllable with accent) for the three speakers (averaged on the eight target words).

Accentual patterns	Speaker	Acoustic properties of the first syllable				Acoustic properties of the second syllable			
		Syllable duration (ms)	f0 min* (Hz)	f0 max* (Hz)	f0 rise (%)	Syllable duration (ms)	f0 min* (Hz)	f0 max* (Hz)	f0 rise (%)
Unaccented word (e.g. ballon)	Male	126	106	104	-2	128	102	96	-6
	Female 1	135	195	191	-2	133	191	181	-5
	Female 2	141	219	210	-4	138	207	197	-5
	<b>Mean</b>	<b>134</b>	<b>173</b>	<b>168</b>	<b>-3</b>	<b>133</b>	<b>167</b>	<b>158</b>	<b>-5</b>
Accented on the first syllable (e.g. BALLon)	Male	261	125	162	30	178	99	94	-4
	Female 1	257	210	301	44	183	182	173	-5
	Female 2	321	224	329	48	187	214	205	-4
	<b>Mean</b>	<b>280</b>	<b>186</b>	<b>264</b>	<b>41</b>	<b>183</b>	<b>165</b>	<b>157</b>	<b>-4</b>
Accented on the second syllable (e.g. baLLON)	Male	174	105	95	-9	246	95	255	169
	Female 1	201	193	173	-10	259	172	442	157
	Female 2	196	206	191	-7	277	190	549	189
	<b>Mean</b>	<b>190</b>	<b>168</b>	<b>153</b>	<b>-9</b>	<b>261</b>	<b>152</b>	<b>415</b>	<b>172</b>

\*For unstressed syllables, the minimum and maximum f0 values correspond to the values associated with the beginning and the end of the f0 plateau.

Statistical analyses performed on the 8 target words produced by the three speakers were then performed on the two acoustic parameters. These analyses showed that for the unaccented words, the length and the f0 rise associated with the first syllables were not different from those of the second syllables [syllable duration:  $t(23)=0.77$ ,  $p > .20$ ; f0 rise:  $t(23)= 1.31$ ;  $p > .20$ ]. On the contrary, for the words that bore accent on their first syllable, the first syllables were longer and associated with a stronger f0 rise than the second syllables [syllable duration:  $t(23) = 7.75$ ,

$p < .0001$ ;  $f_0$  rise:  $t(23)=13.14$ ;  $p < .0001$ ]. Finally, for the words that bore accent on their second syllable, the second syllables were longer and associated with a stronger  $f_0$  rise than the first syllables [syllable duration:  $t(23) = 8.05$ ,  $p < .0001$ ;  $f_0$  rise:  $t(23) = 30.19$ ;  $p < .0001$ ]. The root-mean-squared (RMS) amplitude level for all the target words was equated to 70 dB. Figure 1 illustrates the acoustic properties of the word *ballon* /balɔ̃/ “ball” produced by our male speaker in its unaccented version, with a primary accent on its second syllable and with a primary accent on its first syllable.

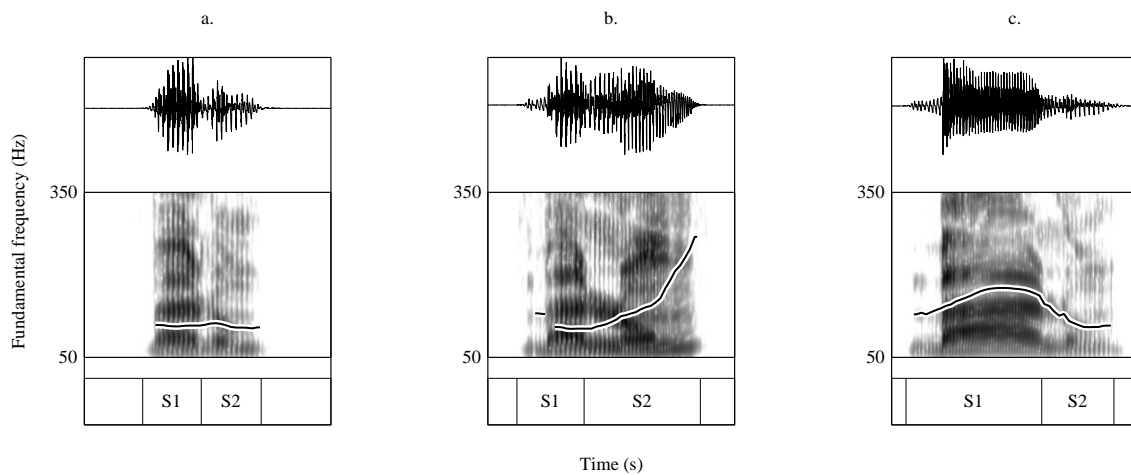


Figure 1. Phonemic and prosodic profile for the word /balɔ̃/ “ball” produced by the male voice in its unaccented version (a), with primary accent on its second syllable (b) and with primary accent on its first syllable (c).

To minimize the involvement, via ipsilateral projections, of the same hemisphere as the ear receiving the stimuli, an audio file containing pink noise was presented to participants in the ear opposite to the one in which the stimuli were presented (see Gonzalez & McLennan, 2007). Pink noise was chosen because as for speech, its spectral level decreases with increasing frequency. Our file containing pink noise was 612 ms long, corresponding to the duration of

the longest target word. RMS amplitude of the signal containing pink noise was equated to 50 dB.

Each trial consisted of three stimuli: A, B and X, with the first two spoken by the two females and the third one by the male. Within each accentual pattern (“native accentual contrast”, “non-native accentual contrast”), A and B varied either in accent, in one phoneme or in both accent and one phoneme, thus leading to 6 experimental conditions (Table 1). The four pairs of words were used in each experimental condition. For each pair of words and within each experimental condition, 16 combinations were used. These combinations resulted from the crossing between accentual pattern (2 possible versions each time: e.g., *ballon* and *baLLON* “ball”), segmental content (2 possible words each time: e.g., *ballon* “ball” and *ballot* “bundle”), response type (2 possible responses: A and B) and type of voice (2 possible voices: female 1 and female 2). 384 trials were thus obtained (6 experimental conditions x 4 pairs of words x 16 combinations).

Procedure: Participants were tested in a sound attenuated booth, and the stimuli were presented over headphones at the same comfortable sound level (60dBA) for all participants. For half the participants, the stimuli were presented to the left ear, while pink noise was presented to the opposite ear. The contrary was true for the other half. Each experimental trial consisted of three stimuli (A, B and X) separated by an interval of 500ms. The participants were told that they would hear stimuli in one ear while hearing pink noise in the opposite ear. They were also told that the first two stimuli would be different, and that the third one would be similar to either the first one or the second one. They were required to press a button on their left or on their right to indicate whether X was identical to respectively A or B. Each participant heard the 384 trials and the experiment lasted approximately 30 minutes. The trials were presented randomly, and

the order of presentation was different for each participant. An inter-trial interval of 1000 ms elapsed between each participant’s response and the beginning of the next trial. The participants began the experiment with a block of 12 practice trials.

**Results and discussion**

Data are available at the Open Resources and Tools for Language (ORTOLANG) repository (<https://hdl.handle.net/11403/accentsdiscrimination>). The percentage of correct responses for the right and the left ear of presentation under each mismatch condition is presented in Figure 2a for the native accentual pattern and in Figure 2b for the non-native accentual pattern.

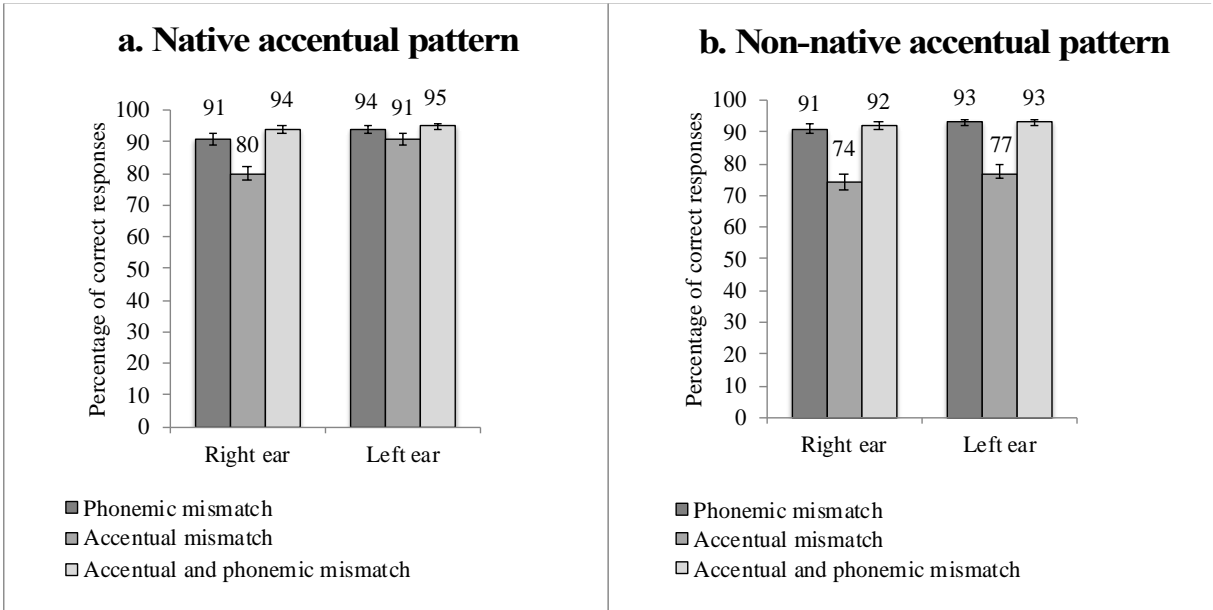


Figure 2. Percentage of correct responses as a function of the ear of presentation (right/left) and the type of mismatch (phonemic/accental/accental and phonemic) for the native accentual pattern (a) and the non-native accentual pattern (b) in Experiment 1. Error bars represent standard error.

We analyzed accuracy data (1=correct response, 0= incorrect response) using a mixed-effect regression model (lme4 package in R-studio statistics Version 1.1.456) with a logistic linking function (Baayen, Davidson & Bates, 2008). The model included mismatch (phonemic, accentual, phonemic and accentual), accentual pattern (native vs. non-native), ear of presentation (right vs. left) and their interactions as fixed effects. The model also included participants and items as random intercepts, plus random participant slopes for the within-participant factors mismatch and accentual pattern, and item slopes for the within-item factors ear of presentation and mismatch (see Barr, Levy, Scheepers & Tily, 2013). Since our data set was too large for the default fitting methods in *glmer.nb*, it was necessary to implement the function option *nAGQ=0* for the model-fitting process to converge (Bates, Maechler, Bolker & Walker, 2015). Significance of variables was assessed using the Anova function from the *car* package (Fox, 2016). The model revealed a significant mismatch x accentual contrast x ear of presentation interaction ( $X^2= 23.02$ ,  $p < .0001$ ). This three-way interaction was examined in two mixed-effect models on native and non-native patterns respectively. The two models included both mismatch and ear of presentation as fixed factors, random intercepts by participant and by item, random participant slopes for the within-participant factor mismatch, and item slopes for the within-item factors ear of presentation and mismatch (Barr et al. 2013). The *nAGQ=0* function option was implemented for both models (Bates et al., 2015). Significance of variables was assessed using the Anova function from the *car* package and multiple comparisons were obtained using the *glht* function from the *multcomp* package with a Bonferonni correction.

For **the native pattern**, the model revealed a significant main effect of the ear of presentation ( $X^2= 6.87$ ,  $p < .01$ ) with more correct responses when stimuli were presented to

the left ear than to the right ear. The effect of mismatch was also significant ( $X^2= 31.46$ ,  $p < .0001$ ) with more correct responses in the phonemic mismatch condition than in the accentual mismatch condition ( $z=3.29$ ,  $p < .01$ ), and more correct responses in the accentual and phonemic mismatch condition than in the accentual mismatch condition ( $z=4.87$ ,  $p < .0001$ ). There was no significant difference between the accentual and phonemic mismatch condition and the phonemic mismatch condition ( $z=1.53$ ,  $p > .20$ ). Crucially, the interaction between the ear of presentation and the mismatch was significant ( $X^2= 10.14$ ,  $p < .01$ ). When stimuli were presented to the left ear, no significant difference was found between the mismatch conditions (all  $p$ .values  $> .20$ ). On the contrary, when stimuli were presented to the right ear, the model revealed more correct responses in the phonemic mismatch condition compared to the accentual mismatch condition ( $z=4.61$ ,  $p < .0001$ ). There were also more correct responses in the accentual and phonemic mismatch condition compared to the accentual mismatch condition ( $z=7.20$ ,  $p < .0001$ ). The difference between the accentual and phonemic mismatch condition and the phonemic mismatch condition was not significant ( $z=0.98$ ,  $p > .20$ ).

For **the non-native pattern**, the model also revealed a significant main effect of mismatch ( $X^2= 109.26$ ,  $p < .0001$ ) with more correct responses in the phonemic mismatch condition than in the accentual mismatch condition ( $z=8.49$ ,  $p < .0001$ ), and more correct responses in the accentual and phonemic mismatch condition than in the accentual mismatch condition ( $z=10.27$ ,  $p < .0001$ ). There was no significant difference between the accentual and phonemic mismatch condition and the phonemic mismatch condition ( $z=0.10$ ,  $p > .20$ ). The effect of ear of presentation was not significant ( $X^2= 0.81$ ,  $p > .20$ ). Crucially the interaction between the ear of presentation and the mismatch was not significant ( $X^2= 0.10$ ,  $p > .20$ ).

To sum up, Experiment 1 showed an effect of the ear of presentation only for the native accentual contrast, with phoneme-like performance when stimuli were presented in the left ear, and thus when processing is constrained to the right hemisphere. In contrast, no effect of the ear of presentation was found for the non-native accentual contrast, and thus whatever the ear of presentation, French participants were still worse at processing accentual differences than phonemic differences. Before discussing more fully the implications of the present results, we conducted another experiment with more controlled stimuli, especially in terms of pitch rise. Indeed, a closer look at the acoustic characteristics of the stimuli used in Experiment 1 revealed that the pitch rise for the stimuli accented on the second syllable was 172% whereas it was of only 41% for the stimuli accented on the first syllable. Thus, the difficulties of French participants with the non-native contrast, especially when stimuli were presented in the left ear, could merely result from the fact that the accentuation on the first syllable was not sufficiently acoustically salient to be perceived. We thus re-ran Experiment 1 with more controlled stimuli on f0.

## **Experiment 2**

Experiment 2 consisted of a replication of Experiment 1 using resynthesized stimuli so that the pitch rise reached that of the stimuli accented on the second syllable. Also, the duration parameter was controlled so that it was the same for the native and non-native accentual contrasts.

## **Method**

Participants: Fifty right-handed French native speakers from the same pool as in Experiment 1

participated in the experiment. Each participant gave informed consent prior to experiment and was remunerated for their time. As in Experiment 1, half the participants heard stimuli in the right ear while the other half heard stimuli in the opposite ear.

Materials: The materials used in Experiment 2 was derived from natural versions of accented (1<sup>st</sup> syllable with accent and 2<sup>nd</sup> syllable with accent) and unaccented words used in Experiment 1. First, the f<sub>0</sub> curve was flattened for all unaccented syllables by adjusting the endpoint of the f<sub>0</sub> plateau to the same value of the starting point of the f<sub>0</sub> plateau. Second, for each word accented on its first syllable, the f<sub>0</sub> maximal value was increased so that the f<sub>0</sub> rise was equal to the f<sub>0</sub> rise of the word accented on its second syllable. For example, before resynthesis, the f<sub>0</sub> maximal value of the word *ballon* accented on its first syllable and produced by the male speaker was 152 Hz, and the pitch rise was 18% (f<sub>0</sub> minimal value: 129Hz). In contrast, the f<sub>0</sub> maximal value of the word *ballon* accented on its second syllable and produced by the male speaker was 305 Hz, and the pitch rise was 190% (f<sub>0</sub> minimal value: 105Hz). As a result, for Experiment 2, the f<sub>0</sub> maximal value of the word *ballon* accented on its first syllable and produced by the male speaker was increased to 374 Hz so that the f<sub>0</sub> rise was now 190% (f<sub>0</sub> minimal value: 129Hz). For the duration parameter, the duration of all unaccented syllables (first and second syllables of unaccented word, second syllable of words accented on their first syllable, first syllable of words accented on their second syllable) were considered and we selected the shortest duration. The duration of all unaccented syllables was then adjusted to this shortest duration. Among the accented syllables (first syllable of words accented on their first syllable and second syllable of words accented on their second syllable), the longest duration was selected, and then the duration of all accented syllables was adjusted to this longest duration. For example, before resynthesis, the duration of the first syllable of the word *ballon* accented on its first syllable and produced by the first female speaker was 356ms while the



duration of the second syllable of the word *ballon* accented on its second syllable was 269ms. For Experiment 2, the duration of the word *ballon* accented on its second syllable was thus increased to 356ms. Hence, the difference in duration between the accented and the unaccented syllable was 224ms for both *ballon* accented on its second syllable and *ballon* accented on its first syllable (i.e., duration of the unaccented syllables: 132ms). All acoustic manipulations were done using the Pitch Synchronous Overlap and Add method (PSOLA, Moulines & Charpentier 1990), which is known for the high degree of naturalness achieved. The acoustic properties of the re-synthesized stimuli are shown in Table 3.

Procedure: It was the same as in Experiment 1.

## **Results and discussion**

Data are available at the Open Resources and Tools for Language (ORTOLANG) repository (<https://hdl.handle.net/11403/accentdiscrimination>). The percentage of correct responses for the right and the left ear of presentation under each mismatch condition is presented in Figure 3a for the native accentual pattern and in Figure 3b for the non-native accentual pattern. Exactly the same analyses as Experiment 1 were performed. A significant mismatch x accentual contrast x ear of presentation interaction was again observed ( $X^2= 12.46$ ,  $p < .01$ ). As in Experiment 1, separate analyses on native and non-native patterns were thus performed.

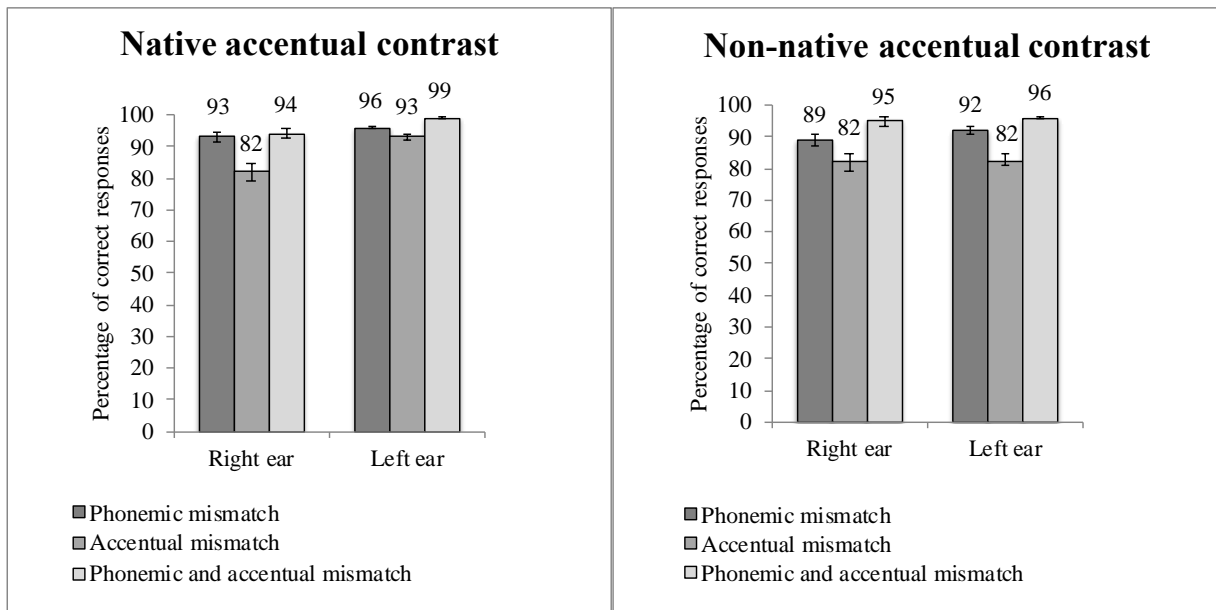


Figure 3. Percentage of correct responses as a function of the ear of presentation (right/left) and the type of mismatch (phonemic/accentual/accentual and phonemic) for the native accentual pattern (a) and the non-native accentual pattern (b) in Experiment 2. Error bars represent standard error.

For **the native pattern**, the model revealed a significant main effect of the ear of presentation ( $X^2= 11.13$ ,  $p < 0.001$ ) with more correct responses when stimuli were presented to the left ear than to the right ear. The effect of mismatch was also significant ( $X^2= 38.66$ ,  $p < .0001$ ) with more correct responses in the phonemic mismatch condition than in the accentual mismatch condition ( $z=3.87$ ,  $p < .001$ ), and more correct responses in the accentual and phonemic mismatch condition than in the accentual mismatch condition ( $z=7.11$ ,  $p < .0001$ ). The difference between the accentual and phonemic mismatch condition and the phonemic mismatch was also significant ( $z=4.63$ ,  $p < .0001$ ). Crucially, the interaction between the ear of presentation and the mismatch was significant ( $X^2= 8.77$ ,  $p < .05$ ). When stimuli were presented to the left ear, no significant difference was found between the phonemic mismatch condition compared to the accentual mismatch condition ( $z=1.90$ ,  $p > .20$ ). The difference between the accentual and phonemic mismatch condition and the phonemic mismatch condition was

significant ( $z=3.98$ ,  $p < .01$ ). Also, the difference between the accentual and phonemic mismatch condition and the accentual condition was significant ( $z=4.45$ ,  $p < .001$ ). On the contrary, when stimuli were presented to the right ear, the model revealed more correct responses in the phonemic mismatch condition compared to the accentual mismatch condition ( $z=3.70$ ,  $p < .01$ ). There were also more correct responses in the accentual and phonemic mismatch condition compared to the accentual mismatch condition ( $z=5.44$ ,  $p < .0001$ ). The difference between the accentual and phonemic mismatch condition and the phonemic mismatch condition was not significant ( $z=1.37$ ,  $p > .20$ )

For **the non-native pattern**, the model also revealed a significant main effect of mismatch ( $X^2= 137.94$ ,  $p < .0001$ ) with more correct responses in the phonemic mismatch condition than in the accentual mismatch condition ( $z=4.41$ ,  $p < .0001$ ), and more correct responses in the accentual and phonemic mismatch condition than in the accentual mismatch condition ( $z=12.58$ ,  $p < .0001$ ). Also, the difference between the accentual and phonemic mismatch condition and the phonemic mismatch condition was significant ( $z=5.41$ ,  $p < .0001$ ). The effect of ear of presentation was not significant ( $X^2= 0.58$ ,  $p > .20$ ). Crucially the interaction between the ear of presentation and the mismatch was not significant ( $X^2= 2.34$ ,  $p > .20$ ).

To sum-up, as in Experiment 1, French participants were worse at processing non-native accentual differences than phonemic differences, and this was the case regardless of whether the stimuli were presented to the right or the left ear. Hence, the difficulties of French listeners with the non-native contrast cannot be merely explained by a lower acoustic salience, especially for pitch rise, for this contrast, in comparison to the native contrast. Note however, that contrary to Experiment, 1 a benefit in the discrimination ability was observed when the difference between words was carried out by accentual difference in addition to phonemic difference for

both the native and the non-native accentual contrasts<sup>4</sup>. It should also be noted that even if French participants showed difficulty in the non-native contrast, their performance was above the chance level (see also Dupoux et al., 1997; Michelas et al., 2018). This boost in the performance in the case of redundant information both for the native and the non-native contrast could be explained by the fact that syllable lengthening was larger for resynthesized stimuli than for natural stimuli.

### **General discussion**

The aim of this study was to examine how the ear of presentation affects the way French listeners use accentual variation to discriminate between words. Because accentual variation at the word level has not a contrastive role in French, our study was built on the strong assumption that this type of variation would be processed as non-linguistic variation by French listeners. Based on McLennan & Gonzalez (2007)'s study showing that non-linguistic information is mainly processed by the right hemisphere, we predicted that French listeners would process an accentual mismatch with the same ability as a phonemic mismatch when processing was pushed into the right hemisphere, and thus when stimuli were presented to the left ear.

Contrary to our predictions, no effect of the ear of presentation was observed for the non-native contrast. Hence, whatever the ear of presentation, French listeners were worse at processing accentual differences than phonemic differences. Such a result replicates Dupoux et al.' studies (Dupoux et al., 1997; Dupoux et al, 2001; Dupoux et al., 2008) in showing that French listeners have persistent difficulty with accentual variation, which does not exist in their

native language. The situation was different for the native accentual contrast. In accordance with our prediction, French listeners performed equally well in the accentual and in the phonemic condition provided that the stimuli were presented to the left ear. Hence, when processing is constrained to the right hemisphere, French listeners are equally good at processing accentual and phonemic differences. In contrast, and perhaps more crucially, French listeners experienced difficulty processing their native accentual contrast when words were presented to the right ear, and thus when processing was constrained to the left hemisphere. Consequently, despite the fact that French listeners routinely produce and perceive words in both their accented and unaccented versions, this native contrast is processed as a non-native contrast when it is managed by the cerebral hemisphere which is dominant for language processing. Finally, the processing of phonemic contrast was not influenced by the ear of presentation, and similar performance was observed when the processing was pushed into the right hemisphere and when it was pushed into the left hemisphere. Such a result thus adds to the growing body of evidence showing that phonemic information is bilaterally represented (Hickok & Poeppel, 2000).

Together, our results strengthen all the studies showing that the phonological system of the native language, and especially here, its prosodic system, influences the way listeners perceive speech sounds (see Best, McRoberts & Goodell, 2001, for a review). Because in French, words cannot receive primary accent anywhere else than on their last syllable, French listeners encounter difficulty discriminating words that differ in stress location. Therefore, as soon as an accentual contrast does not exist in the native language, it is rather hard to overcome the perception difficulty induced by this contrast. In contrast, due to the specific characteristics of French prosody, French listeners are rather good at managing accentual

variation due to the presence vs. absence of an accent, namely the variation that corresponds to the accentual patterns present in their native language (also see Michéas et al., 2018).

Similarly to the right hemisphere advantage found by Gonzalez & McLennan (2007) for the processing of talker-specific variation, we showed a right hemisphere advantage for the processing of accentual variation that exists in French. This right hemisphere advantage could be explained by the fact that accentual information is not linguistically contrastive at the word level for French listeners. Concerning the non-native accentual contrast, the lack of a right hemisphere advantage is likely due to the fact French listeners are not used to process this non-linguistic information. Hence, a training on this type of contrast could be useful to boost the performance especially when the processing is constrained to this right hemisphere. Our claim that accentual variation at the word level is processed as non-linguistic variation by French listeners is in accordance with a previous study focusing on the perception of lexical tones (Hallé, Chang & Best, 2004). In this study, the authors compared the way French and Taiwanese listeners perceived the tonal continua created by manipulating both  $f_0$  and the intensity of Mandarin Chinese tones. Participants performed an ABX task in which they had to decide whether X was identical to either A or B. A and B represented the continuum endpoints and X varied from one endpoint to the other through eight steps. The results showed that Taiwanese listeners had better performance when X was near the two extremities of the continuum than when X was located in the middle of the continuum. In contrast, French listeners had similar performance regardless of the location of X on the continuum. Together, these results showed that French and Taiwanese listeners process tonal information differently. Because tonal information creates meaning distinctions in Mandarin Chinese, it is linguistically relevant for Taiwanese listeners. Consequently, they perceive this information categorically. In contrast, because tonal information is not used to distinguish words in French, it is not linguistically

relevant for French listeners. Consequently, they do not create abstract categories based on this information. Again, these observations mirror those observed for some segmental contrasts that have been shown to be processed at an acoustic level of processing when they do not exist in the native language (Llompert & Reinisch, 2019).

To conclude, our results are in line with Van Lancker (1980)'s hypothesis claiming that prosodic information is differently lateralized in the brain depending on the linguistic functions of prosody in the language. Because accentual information in French is not relevant at the word level, we showed that this information is primarily processed by the right hemisphere. Consequently, our study revealed that a native contrast is processed as a non-native contrast when the processing is constrained to the cerebral hemisphere which is dominant for language processing. Finally, our study also showed that, for French native listeners, a non-native accentual contrast never reaches the performance of a native contrast, and this regardless of the cerebral hemisphere that primarily manages this information.

## Footnotes

<sup>1</sup> Because in French, the acoustic prominence given to a syllable is determined post-lexically and not lexically, the term ‘accent’ rather than ‘stress’ is usually used.

<sup>2</sup> Even though an initial (or secondary) accent can optionally affect the first syllable of French words, this accent does not have the same acoustic properties than the primary accent which is the accent of interest in this study (see Welby, 2006).

<sup>3</sup> Even if the activation dissipates from one cerebral hemisphere to the other via the inter-hemispherical connections and the corpus callosum, manipulating hemifield presentation proved to be effective to assess the respective role of each hemisphere in the processing of variation for both visual (e.g., Marsolek, 2004) and auditory words (see Gonzalez & McLennan, 2007).

<sup>4</sup> Note that the same level of performance was observed in the two experiments. Hence, the fact that we did not observe a significant difference between the phonemic condition and the accentual and phonemic condition in Experiment 1 is likely not due to a ceiling effect.



## References

- Boersma, P., & Weenink, D. (2020). Praat. Doing phonetics by computer (Version 6.1.09, 2020) [Computer program]. Retrieved from <https://www.praat.org>.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390-412.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255-278.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48.
- Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *The Journal of the Acoustical Society of America*, 109(2), 775-794.
- Darcy, I., Daidone, D., & Kojima, C. (2013). Asymmetric lexical access and fuzzy lexical representations in second language learners. *The Mental Lexicon*, 8(3), 372-420.
- Díaz, B., Mitterer, H., Broersma, M., & Sebastián-Gallés, N. (2012). Individual differences in late bilinguals' L2 phonological processes: From acoustic-phonetic analysis to lexical access. *Learning and Individual Differences*, 22(6), 680-689.
- Dupoux, E., Pallier, C., Sebastián, N., & Mehler, J. (1997). A distressing “deafness” in French?. *Journal of Memory and Language*, 36(3), 406-421.
- Dupoux, E., Peperkamp, S., & Sebastián-Gallés, N. (2001). A robust method to study stress “deafness”. *The Journal of the Acoustical Society of America*, 110(3), 1606-1618.
- 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.

- Fox, J. (2016). *Applied Regression Analysis and Generalized Linear Models*, Third Edition. Sage.
- Gandour, J., Wong, D., & Hutchins, D. (1998). Pitch processing in the human brain is influenced by language experience. *Neuroreport*, 9, 2115–2119.
- Gandour, J., Wong, D., Hsieh, L., Weinzapfel, B., Lancker, D. V., & Hutchins, G. D. (2000). A crosslinguistic PET study of tone perception. *Journal of Cognitive Neuroscience*, 12(1), 207-222.
- Gandour, J., Wong, D., Lowe, M., Dziedzic, M., Sathamnuwong, N., Tong, Y., & Li, X. (2002). A cross-linguistic fMRI study of spectral and temporal cues underlying phonological processing. *Journal of Cognitive Neuroscience*, 14(7), 1076-1087.
- Gandour, J., Tong, Y., Wong, D., Talavage, T., Dziedzic, M., Xu, Y., ... & Lowe, M. (2004). Hemispheric roles in the perception of speech prosody. *Neuroimage*, 23(1), 344-357.
- González, J., & McLennan, C. T. (2007). Hemispheric differences in indexical specificity effects in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 33(2), 410.
- Hallé, P. A., Chang, Y. C., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics*, 32(3), 395-421.
- Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences*, 4, 131–138.
- Klein, D., Zatorre, R., Milner, B., & Zhao, V. (2001). A cross-linguistic PET study of tone perception in Mandarin Chinese and English speakers. *Neuroimage*, 13, 646–653.
- Llompart, M., & Reinisch, E. (2019). Imitation in a second language relies on phonological categories but does not reflect the productive usage of difficult sound contrasts. *Language and Speech*, 62(3), 594-622.

- Marsolek, C. J. (2004). Abstractionist versus exemplar-based theories of visual word priming: A subsystems resolution. *The Quarterly Journal of Experimental Psychology, Section A*, 57(7), 1233-1260.
- Michelas, A., & Dufour, S. (2019). Are Prosodic Variants Stored in the French Mental Lexicon?. *Experimental Psychology*, 66 (6), 393-401.
- Michelas, A., Esteve-Gibert, N., & Dufour, S. (2018). On French listeners' ability to use stress during spoken word processing. *Journal of Cognitive Psychology*, 30(2), 198-206.
- Michelas, A., Frauenfelder, U. H., Schön, D., & Dufour, S. (2016). How deaf are French speakers to stress?. *The Journal of the Acoustical Society of America*, 139(3), 1333-1342.
- Moulines, E., & Charpentier, F. (1990). Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Communication*, 9(5-6), 453-467.
- Van Lancker, D. (1980). Cerebral lateralization of pitch cues in the linguistic signal. *Research on Language & Social Interaction*, 13(2), 201-277.
- Van Lancker, D. R., & Canter, G. J. (1982). Impairment of voice and face recognition in patients with hemispheric damage. *Brain and Cognition*, 1(2), 185-195.
- von Kriegstein, K., Eger, E., Kleinschmidt, A., & Giraud, A. L. (2003). Modulation of neural responses to speech by directing attention to voices or verbal content. *Cognitive Brain Research*, 17(1), 48-55.
- Welby, P. (2006). French intonational structure: Evidence from tonal alignment. *Journal of Phonetics*, 34(3), 343-371