Intonation and talker variability in the discrimination of Spanish lexical stress contrasts by Spanish, German and French listeners

Schwab, Sandra ; Dellwo, Volker

DOI: https://doi.org/10.1121/1.5008849

Originally published at:
Schwab, Sandra; Dellwo, Volker (2017). Intonation and talker variability in the discrimination of Spanish lexical stress contrasts by Spanish, German and French listeners. Journal of the Acoustical Society of America, 142(4):2419-2429.
DOI: https://doi.org/10.1121/1.5008849

Intonation and talker variability in the discrimination of Spanish lexical stress contrasts by Spanish, German and French listeners

Sandra Schwab[a] and Volker Dellwo
Institut für Computerlinguistik, Universität Zürich, Andreastrasse 15, 8050 Zürich, Switzerland

(Received 8 November 2016; revised 7 October 2017; accepted 9 October 2017; published online 27 October 2017)

The perception of stress is highly influenced by listeners' native language. In this research, the authors examined the effect of intonation and talker variability (here: phonetic variability) in the discrimination of Spanish lexical stress contrasts by native Spanish ($N = 17$), German ($N = 21$), and French ($N = 27$) listeners. Participants listened to 216 trials containing three Spanish disyllabic words, where one word carried a different lexical stress to the others. The listeners' task was to identify the deviant word in each trial (Odd-One-Out task). The words in the trials were produced by either the same talker or by two different talkers, and carried the same or varying intonation patterns. The German listeners' performance was lower compared to the Spanish listeners but higher than that of the French listeners. French listeners performed above chance level with and without talker variability, and performed at chance level when intonation variability was introduced. Results are discussed in the context of the stress "deafness" hypothesis.
 https://doi.org/10.1121/1.5008849

[TCB] Pages: 2419–2429

I. INTRODUCTION

Native speakers of a language without lexical stress contrasts are disadvantaged in perceiving such contrasts in foreign languages that have stress contrasts. This phenomenon is often referred to as stress "deafness" (e.g., Dupoux *et al.*, 1997, 2001, 2008). According to the stress deafness hypothesis, the perception of lexical stress is strongly affected by whether lexical stress is free or fixed in a listener's native language. In free-stress languages such as Spanish, German, or English the position of stress, which is realized at the word level (hence the name *lexical stress*), is defined by morphophonological constraints, and varies between words. Words in free-stress languages can have stress on the antepenultimate syllable (e.g., es. pájaro, de. Einrichtung, en. critical)[1], on the penultimate syllable (e.g., es. rosa, de. Arbeit, en. syllable), or on the final syllable (e.g., es. correr; de. Kontakt, en. parade). In such languages, lexical stress can be contrastive (e.g., es. *número* vs *numero*, the number vs I number; de. *umfahren* vs *umfahren*, run into vs drive round; en. *record* vs *record*). Although the assignment of lexical stress is still a controversial issue, there is a common agreement on the fact that stress patterns, in free-stress languages, are somehow involved in the lexical representations. However, it is unclear whether the stress patterns of all lexical items or only irregular stress patterns are stored, or whether the stress assignment combines the retrieval of a stored representation and the computation of stress patterns on the basis of phonological and/or morphological rules or statistical distribution (Laganaro *et al.*, 2002; Levelt *et al.*, 1999; Protopapas *et al.*, 2016). Contrary to free-stress languages, the position of lexical stress in fixed-stress languages (e.g., French, Polish, Turkish, Hungarian, etc.) cannot be contrastive. Consequently, stress information does not need to be stored in lexical representations and can be assigned by default.

So far, it has been demonstrated that listeners of fixed-stress languages experience difficulties in distinguishing stress contrasts in a free-stress foreign language (Dupoux *et al.*, 2008; Peperkamp *et al.*, 2010; Rahmani *et al.*, 2015). It is unclear, however, whether—and if so, to what degree—listeners of a free-stress $L1$ are disadvantaged in a free-stress foreign language where the phonetic realization of stress might be different from their native language. On the one hand, Dupoux *et al.* (1997, 2001, 2008) showed that Spanish listeners had no difficulties in judging stress in Spanish stimuli produced by Dutch speakers. For example, Dupoux *et al.* (1997) reported that the Spanish listeners' error rate in an ABX task (involving stress contrasts) was only 4% (see Table I in Dupoux *et al.*, 1997). On the other hand, Ortega-Llebaria *et al.* (2013) showed that native listeners of (free stress) English were disadvantaged compared to (free-stress) Spanish listeners in judging stress in Spanish stimuli. They showed, in an identification task, that English listeners experienced difficulties perceiving stimuli with realizations of stress that are common in Spanish but not in English (e.g., small duration difference between stressed and unstressed syllables; absence of vowel reduction).

Further, in the Dupoux *et al.* experiments there was no control group (i.e., no Dutch listeners) and the Ortega-Llebaria *et al.* study did not include a group of listeners with a fixed-stress native language. That means that so far there has been no design in which stress perception performance in a free-stress language was directly compared between non-native listeners of a fixed- and a free-stress language in

[a]Also at: Ecole de langue et de civilisation françaises, Université de Genève, Candolle 5, 1211 Genève 4, Switzerland. Electronic mail: sandra.schwab@uzh.ch

comparison to a native control group. This comparison was achieved in the present study by testing the ability of free-stress German, fixed-stress French, and free-stress (native) Spanish listeners to perceive stress contrasts in Spanish.

Acoustic-phonetic variability, for example the type of variability introduced by different voices or different variations in fundamental frequency (*F*0), has a strong influence on stress perception. With the exception of Tremblay (2009), who investigated the effect of talker variability on the perception of word stress by French Canadian learners of English, researchers have not directly addressed the effect of phonetic variability on stress perception in non-native listeners with and without stress deafness. Instead, they have used stimuli with and without phonetic variability in different experiments and then compared the results (e.g., Dupoux *et al.*, 1997, 2001). Further, these studies only assessed one type of phonetic variability, namely, talker variability: variability was introduced either by using two or three talkers or by manipulating *F*0 to simulate different talkers. However, there are other sources of phonetic variability that are relevant to the study of lexical stress contrasts. One of them is phonetic variability introduced by intonation contrasts. Intonation contours enable listeners to distinguish, amongst other things, statements from yes/no questions. Intonation plays an even more important role in languages where yes/no questions might have the same syntax as statements. (e.g., fr. *Tu viens* vs *Tu viens*?, es. *Vienes* vs *Vienes*?, You come vs Do you come?). Yes/no questions in these languages are mainly expressed by means of *F*0 rises, which means that the listeners cannot rely on *F*0 to identify the stressed syllable in the same way as in statements.

In the present experiment, we tested the effects of intonation and talker variability on the discrimination of Spanish stress contrasts by Spanish, German, and French listeners. Intonation variability was introduced by taking words in phrase final position that either stem from a declarative or interrogative (yes/no question) clause. In this context, the use of *F*0 as an intonational cue to grammatical function might compete with its use as a cue to lexical stress. Such intonation variability introduces a level of complexity in the acoustic cues to stress which might be challenging even for native listeners, particularly when words are taken out of context. As for talker variability, stimuli from multiple talkers increase within-category variability in the acoustic cues which mark lexical stress. Talker variability was introduced by presenting listeners with stimuli recorded by two female speakers. The examination of these two sources of phonetic variability enables us to determine to what extent each of them impedes the processing of lexical stress in listeners with and without stress contrasts in their native language.

A. Lexical stress and intonation in Spanish, German, and French

Spanish, German, and French were chosen for the present study because of their accentual and intonational characteristics. As previously mentioned, Spanish and German are free-stress languages where primary stress is realized at the lexical level. In contrast, French is a fixed-stress language in which primary stress is realized at the phrasal level, more precisely on the last syllable of the accentual phrase (Lacheret-Dujour and Beaugendre, 1999). As a consequence, primary stress is contrastive in Spanish and German (e.g., es. *válido* vs *valido*, I validate vs he validated), whereas it has a demarcative function in French (i.e., it marks the right end of the accentual phrase, e.g., Hayes, 1995; Hyman, 2006; Lacheret-Dujour and Beaugendre, 1999). In French, moreover, both the realization of primary stress and the intonation contours are anchored on the final syllable of the accentual phrase, which leads to "fusion" between accentual and intonational structures (i.e., the so-called "syncretism" between accentuation and intonation; Rossi, 1979; Vaissière, 2002).

Further, the acoustic correlates of primary stress—variations in duration, fundamental frequency (*F*0), and intensity—are used in a different way in Spanish, German, and French. Although the phonetic realization of German word stress involves a change in *F*0 and, to a lesser extent, in intensity, the duration of the vowel is considered as the primary cue to German lexical stress (Dogil and Williams, 1999; Jessen *et al.*, 1995). In Spanish, the distinction between stressed and unstressed syllables is expressed by a joint variation of duration and *F*0, and to a lesser degree, by changes in intensity (Quilis, 1981; Llisterri *et al.*, 2014). The main feature that distinguishes German and Spanish word stress is the degree of phonological vowel reduction (Delattre, 1969). In German, vowel reduction can lead to a complete neutralization of some vowel contrasts due to the shortening of unstressed vowels (especially tense vowels) and to formants modifications (i.e., vowel centralization) (Mooshammer and Geng, 2008), whereas vowels never lose their quality in Spanish (Ortega-Llebaria and Prieto, 2009; Quilis, 1981). In French, it has been shown that duration plays the most determinant role in signaling stressed syllables, while intensity is a less important parameter (Delattre, 1938; Léon, 2007). French stressed syllables are also often superimposed with *F*0 rises or *F*0 peaks that are related to prosodic boundaries (on account of the aforementioned syncretism between accentuation and intonation; Rossi, 1979; Vaissière, 2002). The superposition of *F*0 onto stress can also be found in German and Spanish. The main difference, however, is that in French, *F*0 movements come from the marking of prosodic boundaries, while in German and Spanish they come from phase-level accentuation (i.e., different pitch accent types).

Despite the considerable variation found in the intonation contours of neutral statements and yes/no questions (i.e., the two grammatical constructions used in the present research) in Spanish, German, and French (e.g., Hualde and Prieto, 2015; Santiago Vargas and Delais-Roussarie, 2012; Wochner *et al.*, 2015), the following general tendencies are observed. Spanish neutral statements can present a low pitch accent (L*), or under some circumstances (narrow focus, emphasis), a rising pitch accent (L + H*), and final falling pitch contour (L%) (Hualde and Prieto, 2015). In German, neutral statements also present a rising nuclear pitch accent (i.e., L + H*) and a final falling pitch contour (i.e., L%) (Baumann *et al.*, 2000). In contrast, French neutral statements are characterized by a low nuclear pitch accent (i.e., L*) and a final falling pitch contour (i.e., L%) (Delais-Roussarie *et al.*, 2015). Spanish and German yes/no questions are expressed with a low nuclear pitch accent tone (i.e., L*) and a final rising pitch contour (i.e., H%) (Baumann *et al.*, 2000; Hualde and Prieto, 2010). However, a rising nuclear pitch accent tone (i.e., L + H*) can also be observed, at least in Spanish in some particular circumstances (e.g., laboratory speech) (Hualde and Prieto, 2015). French differs again from Spanish and German in this regard, since yes/no questions are marked by a high nuclear pitch accent tone (i.e., H*) and a final rising pitch contour (i.e., H%) (Delais-Roussarie *et al.*, 2015).

In summary, with the exception of vowel reduction, German shares many accentual and intonational similarities with Spanish. On the contrary, French differs from both languages, at the stress as well at the intonation level.

B. Hypotheses and predictions

We examined the ability of German native listeners to detect stress in Spanish, in comparison with the performance of French and Spanish native listeners. According to the stress deafness hypothesis (Dupoux *et al.*, 1997, 2001, 2008)—in its more conservative form—German listeners should show similar performance to native Spanish listeners in detecting stress in Spanish, since both languages have free contrastive stress. However, since the degree of vowel reduction is much less in Spanish than in German (Delattre, 1969), German listeners, who can use vowel quality as a cue to lexical stress in German, cannot use it in Spanish. We thus hypothesized that the different mechanisms for realizing stress in German and Spanish leads to poorer stress detection performance for German compared to Spanish listeners [as shown by Ortega-Llebaria *et al.* (2013) for English listeners], despite the fact that both languages have free stress. Given that stress is fixed in French, French listeners, for their part, were expected to show significantly lower performance than Spanish and German listeners.

As for intonation variability, previous studies have shown that native listeners re-weight these cues relative to lexical stress, taking into account the intonation pattern of the utterance (e.g. Muñoz García, 2010; Ortega-Llebaria, 2009; Ortega-Llebaria and Prieto, 2009). Given these conclusions, we predicted that variability in intonation would not affect the performance of native Spanish listeners. Since German listeners are used to stress encoding mechanisms and since the intonation contours in yes/no questions are very similar in German and Spanish, their performance was not expected to be hampered by the introduction of intonation variability. By contrast, Muñoz García (2010) showed that a rising intonation contour in Spanish lowered the performance of French listeners with beginner-intermediate competence in Spanish. As a consequence, we predicted that the performance of French listeners, who are not able to encode stress representation and whose native language differs considerably from Spanish at the stress level as well as at the intonation level, would be strongly lowered by intonation variability.

Talker variability has been shown to trigger a processing cost that decreases the efficiency of speech perception in native and foreign languages (e.g. Antoniou *et al.*, 2015; Johnson, 2005; Mullennix *et al.*, 1989; Pisoni, 1997; van Dommelen and Hazan, 2012). Nevertheless, Dupoux *et al.* (2001) found that the Spanish listeners' error rate did not increase with the introduction of talker variability in Spanish stimuli (19% without talker variability versus 20% with talker variability), whereas French listeners' error rates almost doubled (27% without talker variability versus 53% with talker variability). Following these results, we expected that Spanish listeners' stress perception would not be affected by talker variability. We hypothesized that German listeners would not be influenced by talker variability, since they are able to encode stress representations. In contrast, we predicted that the performance of French listeners would considerably decrease with the introduction of talker variability.

II. METHOD

A. Participants

Three groups of listeners were tested: Group 1: 17 native speakers of Spanish/Catalan[2] from Catalonia (mean age: 23.4 yrs, stdev: 5.8), Group 2: 21 native German/Swiss German listeners from Switzerland (mean age: 24.5 yrs, stdev: 3.4), Group 3: 27 native French listeners from Switzerland and France (mean age: 22.7 yrs, stdev: 4.2). The native Spanish listeners were students from the Autonomous University of Barcelona and the University Pompeu Fabra (Barcelona), the native German participants were students from the University of Zurich, and the native French participants were students from the University of Neuchâtel. The native German and French listeners had no knowledge of Spanish, Italian, or Portuguese (i.e., free-stress romance languages). Since French, German, and English are mandatory disciplines in the Swiss educational system, the German listeners had knowledge of English and French, and French listeners had knowledge of German and English, but none were bilingual.

B. Material

1. Stimuli and design

Trisyllabic Spanish words which only differed in lexical stress—words with stress on the first syllable (hereafter "1st"), on the second syllable (hereafter "2nd"), and on the third syllable (hereafter "3rd") – were used: *número* (number), *numero* (I number), *numeró* (he numbered); *válido* (valid), *valido* (I validate), *validó* (he validated). Two native female talkers of Peninsular Spanish (Talker 1 and Talker 2) were recruited to record the materials in the Phonetics Laboratory in Zurich. They produced the six words twice in a declarative sentence (12 declarative sentences, 2 talkers × 6 words) with a falling intonation (i.e., es. Le dijo a Pat'número'; en. He/she said to Pat "number") and twice in an interrogative sentence (12 interrogative sentences, 2 talkers x 6 words) with a rising intonation (i.e., es. ¿Le dijo a Pat'número'?; en. Did he/she say to Pat number?). We isolated the final trisyllabic words (from the total 12 declarative + 12 interrogative sentences) which served as stimuli in the perception experiment.

We created trials containing three segmentally identical words separated by 500 ms. Two of the words had the same stress pattern ("standards") and one of them a different one ("deviant"), e.g. deviant 'número' (first-syllable stressed) among standards "numero" (second-syllable stressed). Among all the stimuli (6 words × 2 talkers × 2 intonation patterns = 24 stimuli), only the 6 stimuli produced by Talker 1 with the falling intonation were used as target (i.e., deviant) stimuli in the test trials, while the other stimuli were used to introduce variability within the trial. We constructed 72 trials with the word numero and 72 trials with the word "valido" ($N = 144$) according to the following principles:[3]

(a) The stress pattern of the target stimulus was balanced, i.e., the same number ($N = 24$) of first-, second-, and third-syllable stressed words were used as target stimuli.
(b) All stress contrasts were tested, i.e., 12 first-syllable stressed deviants among second-syllable stressed standards (i.e., "1st-1st-2nd") and 12 among third-syllable stressed standards (i.e., "1st-1st-3rd"); 12 second-syllable stressed deviants among first-syllable stressed standards (i.e., "2nd-2nd-1st") and 12 among third-syllable stressed standards (i.e., "2nd-2nd-3rd"); 12 third-syllable stressed deviants among first-syllable stressed standards (i.e., "3rd-3rd-1st") and 12 among second-syllable stressed standards (i.e., "3rd-3rd-2nd").
(c) The position of the target within the trial (Position 1, Position 2, Position 3) was balanced, i.e., each target stimulus appeared 24 times in each position.
(d) Four experimental conditions were created in equal proportions. In the first condition (*1into-1talker*), the three stimuli within the trial were produced with the same (i.e., falling) intonation pattern by one talker (i.e., Talker 1). In the second condition (*1into-2talkers*), the three stimuli were produced with the same (i.e., falling) intonation pattern by two talkers (Talker 1 and Talker 2). In the third condition (*2into-1talker*), the three stimuli were produced with falling and rising intonation contours by one talker (i.e., Talker 1). Finally, in the fourth condition (*2into-2talkers*), the three stimuli were produced with falling and rising intonation contours by two talkers (i.e., Talker 1 and Talker 2). In each condition, the target stimulus was produced by Talker 1 with falling intonation, which allows us to compare the detection of the same target stimulus across the four experimental conditions (i.e., *1into-1talker*, *1into-2talkers*, *2into-1talker and 2into-2talkers*).

Besides the 144 test trials (2 lexical items [numero, valido] × 6 stress contrasts × 3 positions × 4 conditions), we also introduced 72 filler trials to safeguard against the possibility that participants would develop a strategy to identify the deviant element of the trial. In the filler trials, the target stimulus was produced by Talker 1 with a rising intonation contour ($N = 36$), or by Talker 2 with a falling ($N = 18$), or with a rising intonation ($N = 18$). The filler trials were not included in the analyses.

2. Acoustic description and perceptual evaluations of the test stimuli

We performed an acoustic analysis of the words used in the test trials. These words were first-, second-, and third-syllable stressed words produced by Talker 1 with falling and rising intonation contours and produced by Talker 2 in falling intonation (the words produced by Talker 2 with rising intonation contour only were used in the filler trials and thus were not analyzed). For the three syllables of each word, we measured the duration (in ms), the mean value of $F0$ (in Hz; Hirst, 2011), and the maximum value of intensity (in dB).[4] Both talkers revealed similar patterns in declarative stimuli, independently of the stress pattern, which suggests that comparable acoustic patterns governed the realization of stress. As far as the declarative stimuli are concerned (see Fig. 1, first column), the comparison of the stressed syllable in a particular pattern with the same unstressed syllables in the other patterns (e.g., the syllable "nu" is stressed in the first-syllable stressed word numero, but it is unstressed in the second-syllable stressed numero and in the third-syllable stressed word numero) shows that stress is marked by an increase in duration, especially in first- and second-syllable stressed words (Fig. 1, the two first rows). The increase of the duration to cue stress in the third-syllable stressed word is less noticeable, as it is coupled with the lengthening of the last syllable due to the sentence-final position of the word. Moreover, the comparison of the three stress patterns indicates that the presence of stress is also signaled by increases in both $F0$ and intensity, especially in third-syllable stressed words.

For the interrogative stimuli (see Fig. 1, second column), an increase of duration also indicates the presence of stress, although to a lesser extent than for the declarative stimuli. Again, we observe the lengthening of the final prepausal syllable in the three stress patterns, which reduces the use of duration to signal stress in third-syllable stressed words. As for $F0$, although we do not find an increase of $F0$ on the stressed syllable, we observe a different slope in the $F0$ rise across the three stress patterns. This difference can be explained by the fact that the $F0$ rise begins on the stressed syllable and ends at the end of the word (which is in agreement with Hualde and Prieto, 2015). Finally, an increase in intensity is also used to signal stress, again especially in words with stress on the final syllable. No statistical analyses were performed to test these observations due to the small number of items.

To summarize, for the declarative stimuli, an increase of duration and $F0$ constitutes the main cues to stress in first- and second-syllable stressed words, whereas an increase of $F0$ and intensity signals stress in third-syllable stressed words. In the interrogative stimuli, while duration and intensity are used in the same way as in declarative stimuli, $F0$ does not increase on the stressed syllable, but presents a different rising slope according to the stress pattern of the word.

To ensure the naturalness of the test stimuli, we performed two perceptual evaluations.[5] In the first evaluation, an expert in Spanish prosody annotated the stimuli using

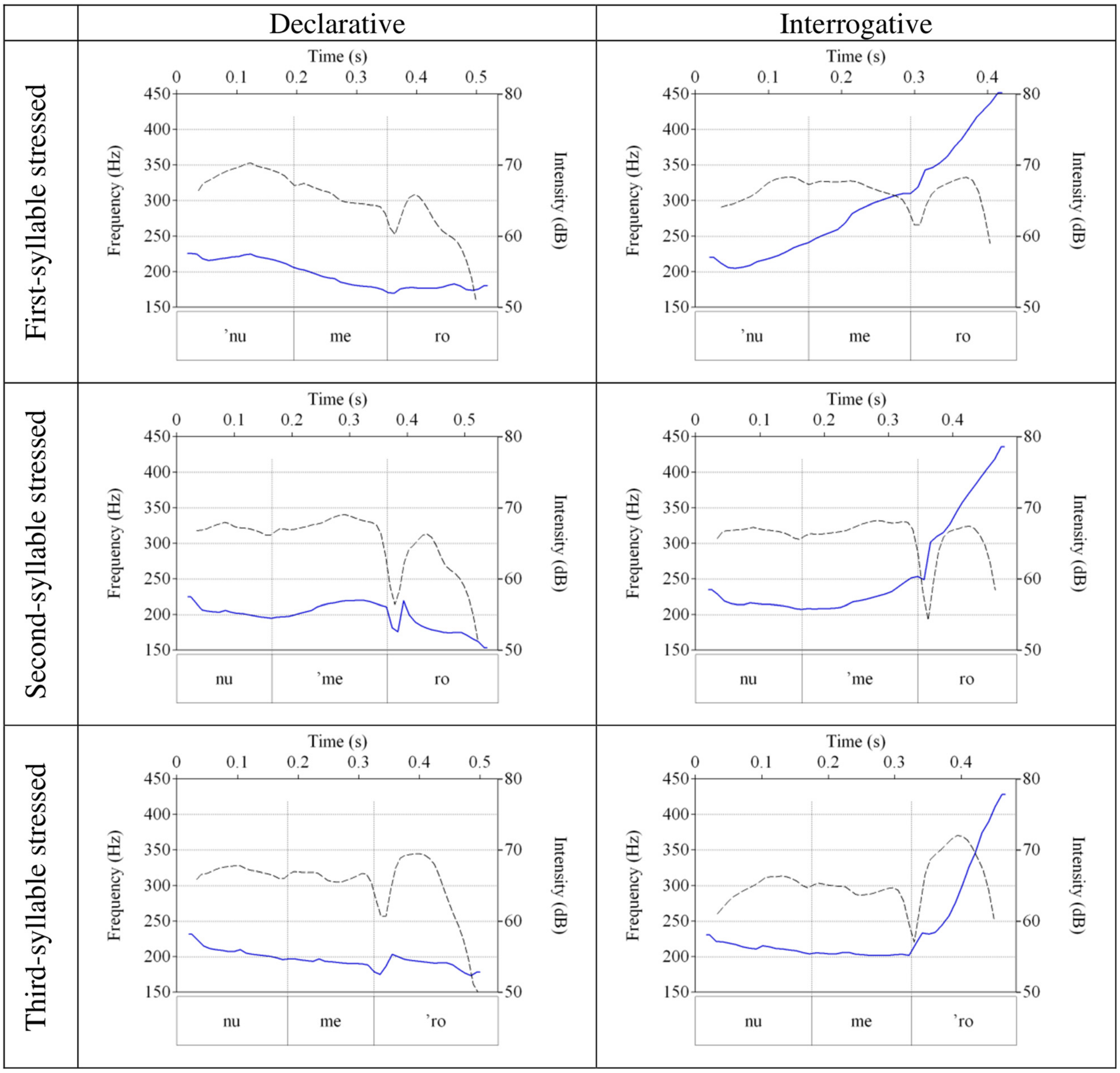

FIG. 1. (Color online) Duration (in seconds; indicated by the syllabic interval), $F0$ (in Hz; solid curve), and intensity (in dB; dashed curve) for the lexical item numero produced by Talker 1 as a function of intonation (columns) and stress patterns (rows).

Sp_ToBi transcription. All declarative stimuli were perceived with a high or rising nuclear pitch accent (i.e., H* or L + H*) and a final falling pitch contour (i.e., L%). Interrogative stimuli were judged as having a rising nuclear pitch accent tone (i.e., L + H*) and a final rising pitch contour (i.e., H%). These results are in line with the nuclear melodic configurations reported in the literature on Spanish statements and yes/no questions. Note that, in these contexts, falling/low nuclear pitch accents (i.e., L* or *) are commonly seen as the unmarked forms (Estevas-Vilaplana and Prieto, 2010; Quilis, 1981). However, high/rising nuclear pitch accents are often observed, especially when data are elicited in laboratory conditions and the speakers mark a certain emphasis on the words they read (Hualde and Prieto, 2015). The second evaluation was performed by ten native Spanish phoneticians. They were instructed to listen to each stimulus and to indicate the stressed syllable. The results of this evaluation showed that the stress pattern of all stimuli was correctly identified by seven or more phoneticians. The two words that received the lowest score (7/10) were the third-syllable stressed words produced by Talker 1 with rising intonation.

C. Procedure

The participants performed an Odd-One-Out task (Schwab and Dellwo, 2016). After hearing each trial, listeners indicated which of the three elements was the deviant element (i.e., "odd-one-out") by pressing the corresponding key (1, 2, or 3) on a response box. Listeners were asked to give their answers as quickly as possible. They were told that the odd element differed with respect to the stressed syllable and were given examples from Spanish stress contrasts to illustrate the task. If participants failed to respond within

3 s, a missing response was recorded and the next trial was presented. Trials were randomized for each listener. The experiment lasted about 25 min for each listener.

D. Data analysis

Statistical analyses were carried out using R (version 3.1.3; R Development Core Team, 2016; lmerTest R package; Kuznetsova *et al.*, 2014). We ran a mixed-effects logistic regression model on the correct/incorrect responses (Baayen *et al.*, 2008). The fixed part of the model was comprised of "language" (Spanish, German, French), "phonetic variability" (1into-1talker, 1into-2talkers, 2into-1talker, 2into-2talkers) and the interaction between language and phonetic variability. We included in the model the following control variables, as well as their interaction with language: "trial number," "lexical item" (valido, numero), "odd position" (1, 2, 3), and "stress contrast" (1st–2nd, 1st–3rd, 2nd–1st, 2nd–3rd, 3rd–1st, 3rd–2nd). The nominal variables were recoded into dummy variables: language (Spanish, German, French) to [0, 1], phonetic variability (1into-1talker, 1into-2talkers, 2into-1talker, 2into-2talkers) to [0, 1], lexical item (valido, numero) to [1, −1], deviant position (1, 2, 3) to [0, 1], and stress contrast (1st–2nd, 1st–3rd, 2nd–1st, 2nd–3rd, 3rd–1st, 3rd–2nd) to [0, 1]. The numerical variable trial number was centered on the mean (for R operational reasons the variable was first rescaled between 0 and 1). The random part of the model included random intercepts for participants and items, and random slopes allowing for the effect of phonetic variability to differ across participants and the effect of language to vary across items. The significance of the main effects and interactions was assessed with likelihood ratio tests that compared the model with the main effect or interaction to a model without it. The estimates (β) are expressed in logit and were computed taking "incorrect response" as the reference level for the dependent variable. *Post hoc* analyses with Tukey correction for multiple comparisons were performed to obtain 2×2 comparisons. The detection of influential data (i.e., data that unduly influences the result of the analysis) was performed using the methodology described in Nieuwenhuis *et al.* (2012; influence.ME R package). Based on Cook's distance and testing for changes in statistical significance (sigtest), no influential data were observed in any of the analyses. The figures show the percentage of correct responses; all statistical analyses were performed on raw data (correct/incorrect responses).

III. RESULTS AND DISCUSSION

Results showed a main effect of language (see Fig. 2) which was highly significant [$\chi^2(2) = 101.02$, $p < 0.001$].[6] *Post hoc* tests showed that the French listeners performed worse (44.62%), compared to the Spanish listeners (90.86%; $\beta = -3.562$, standard error (SE) $= 0.285$, $z = -12.500$, $p < 0.001$) and to the German listeners (76.29%; $\beta = -1.988$, SE $= 0.241$, $z = -8.232$, $p < 0.001$). The German listeners' performance was significantly lower than that of the Spanish listeners ($\beta = -1.574$, SE $= 0.299$, $z = -5.270$, $p < 0.001$). A chi-square analysis showed that although the global rate of missing data was relatively low (4.11%), the French listeners

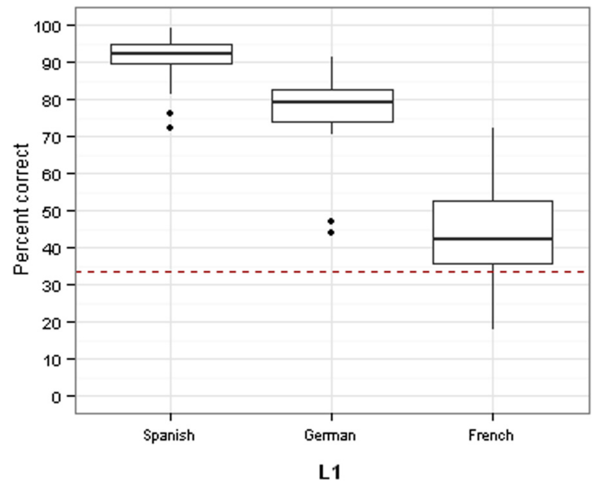

FIG. 2. (Color online) Percent correct identification of the deviant as a function of language (Spanish, German, and French). The dash line represents chance level (33%).

presented more missing data (6.33%, stdev: 5.73) than the German listeners [3.14%, stdev: 4.22; $\chi^2(1) = 36.81$, $p < 0.001$], who in turn presented more missing data than the Spanish listeners [1.8%, stdev: 2.7; $\chi^2(1) = 9.87$, $p = 0.002$].

We also observed a main effect of phonetic variability [$\chi^2(3) = 46.71$, $p < 0.001$] modulated by the effect of language [$\chi^2(6) = 15.27$, $p = 0.018$]. As can be seen in Fig. 3, phonetic variability did not show a similar effect across the three listener groups.[7]

For the Spanish listeners (see Fig. 3, first row), the introduction of talker variability did not significantly affect detection of the deviant stress pattern. This is observed in trials without intonation variability (1into-1talker = 97.18%, 1into-2talkers = 96.21%), as well as in trials with intonation variability (2into-1talker = 87.67%, 2into-2talkers = 81.99%). However, Spanish listeners' performance significantly decreased with the introduction of intonation variability, regardless of whether the trials were produced by one talker (1into-1talker = 97.18%, 2into-1talker = 87.67%) or by two talkers (1into-2talkers = 96.21%, 2into-2talkers = 81.99%). Moreover, the presence of intonation variability without talker variability (i.e., 2into-1talker) led to a poorer performance than the presence of talker variability without intonation variability (i.e., 1into-2talkers).

Although the German listeners performed less well than the native Spanish listeners, they showed the same pattern as the Spanish listeners (see Fig. 3, second row). Independent of intonation variability, their detection of the stress deviant was not influenced by the presence of talker variability (without intonation variability: 1into-1talker = 87.04%, 1into-2talkers = 82.92%; with intonation variability: 2into-1talker = 68.77%, 2into-2talkers = 66.94%). The detection of the stress deviant was, however, less accurate in trials with intonation variability than in trials without intonation variability, whether the trials contained talker variability (1into-2talkers = 82.92%, 2into-2talkers = 66.94%) or not (1into-1talker = 87.04%, 2into-1talker = 68.77%). In line with the Spanish listeners, we observed that the introduction of intonation variability without talker variability (i.e., 2into-1talker) led to a

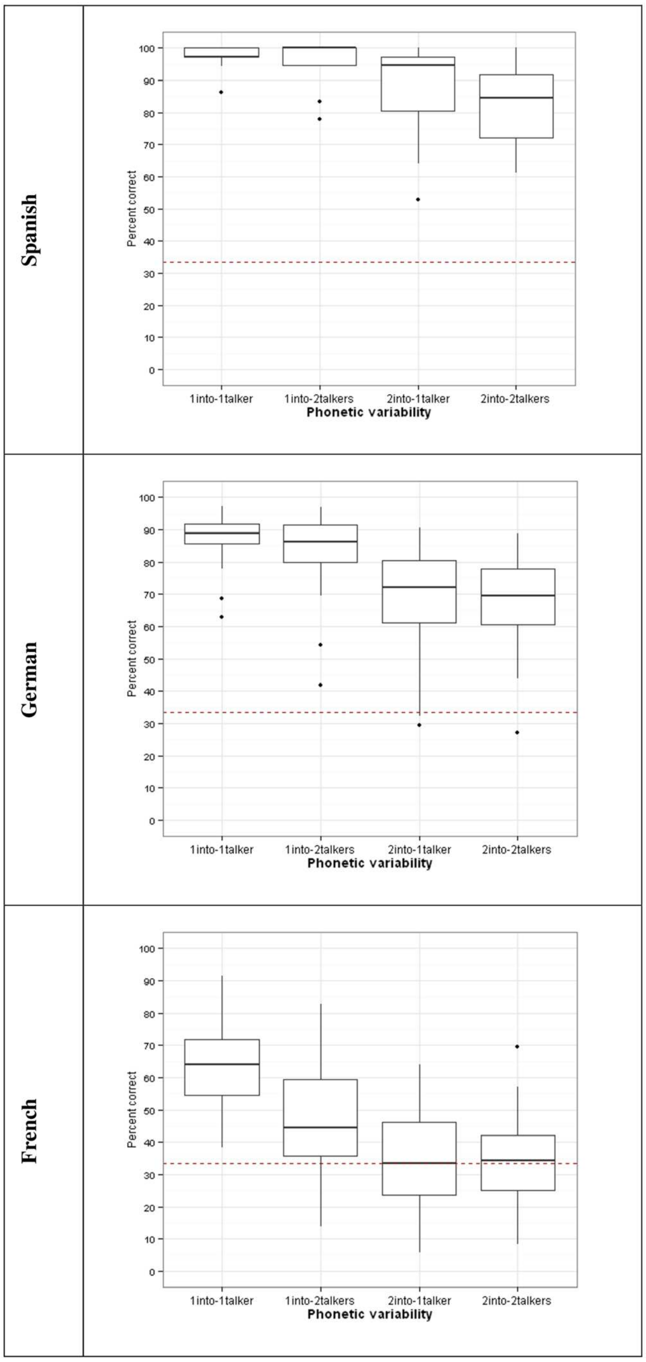

FIG. 3. (Color online) Percent correct identification of the deviant as a function of phonetic variability for the Spanish, German, and French listeners. The dash line represents chance level (33%).

poorer performance than the introduction of talker variability without intonation variability (i.e., 1into-2talkers).

In addition to their lower performance, the French listeners presented a different pattern to the Spanish and German listeners (see Fig. 3, third row). The introduction of talker variability in trials without intonation variability significantly hampered their detection of the stress deviant (1into-1talker = 62.92%, 1into-2talkers = 47.46%). However, the effect of talker variability was not present in trials which included intonation variability (2into-1talker = 35.11%, 2into-2talkers = 33.33%). Similar to the German and Spanish listeners, the presence of intonation variability affected the French listeners' performance, whether the trials were produced by one talker (1into-1talker = 62.92%, 2into-1talker = 35.11%) or by two (1into-2talkers = 47.46%, 2into-2talkers = 33.33%). We also observed that the introduction of intonation variability without talker variability (i.e., 2into-1talker) led to a poorer performance than for the introduction of talker variability without intonation variability (i.e., 1into-2talkers).

Post hoc analyses (with Tukey adjustments) also showed that, for each condition of phonetic variability, the differences between each of the three listener groups were significant ($p < 0.001$). Finally, the Spanish and German listeners' performance was well above chance level (exact binomial p [one-tailed] < 0.05) in the four conditions of phonetic variability. The French listeners' performance was above chance level in their detection of stress deviant when there was no intonation variability (1into-1talker: exact binomial p [one-tailed] < 0.05, 1into-2talkers: exact binomial p [one-tailed] < 0.05), but their performance was at chance level when the stimuli included intonation variability (2into-1talker: exact binomial p [one-tailed] $= 0.14$, 2into-2talkers: exact binomial p [one-tailed] $= 0.51$).

IV. GENERAL DISCUSSION

This research aimed to determine the role of phonetic variability (here: talker and intonation variability) in the perception of Spanish lexical stress by native Spanish, German, and French listeners. Native German listeners, who were not expected to experience difficulties in the identification of the stress deviant—according to the stress deafness hypothesis—performed significantly less well than native Spanish listeners overall (average accuracy of 76% and 90%, respectively). This result means that, despite familiarity with stress contrasts from their own language, they are less successful at detecting lexical stress in Spanish than Spanish listeners. This finding is interesting in regard to the stress deafness hypothesis which—in its most conservative form—would not predict differences between listeners of free-stress languages (e.g., Dupoux *et al.*, 1997), all the more so given the prosodic (i.e., stress and intonation) similarities between both languages. There are three explanations for the German listeners' lower performance. First, although both German and Spanish have words with stress on the antepenultimate, penultimate, and final syllable (e.g., Wiese, 1996), the German stress pattern seems to be related to the syllabic structure of the word. Janssen (2003), for example, examined a corpus of trisyllabic German words ending with different syllabic structures and showed that 58.3% of the words ending with a vowel (i.e., the syllabic structure of the words used in the present experiment) had stress on the penultimate syllable, 37.7% had stress on the antepenultimate, and only 4% had stress on the final syllable. Thus, the native German listeners might have been confused by final stress on a word ending with a vowel, and therefore had difficulty identifying the final syllable as being stressed. A closer examination of the differences between the "stress contrasts" revealed that German listeners experienced more difficulty when first- and third-syllable stressed stimuli were presented together in the same trials. This result would suggest that fine phonological—or rather phonotactic—details can influence listeners' ability to process lexical stress in an unknown language.

Second, as mentioned in Sec. I, German vowels are often reduced in unstressed syllables, which is not the case for Spanish vowels. As a consequence, the German listeners lacked an important cue when they had to identify stress position in Spanish words. This finding is in line with Ortega-Llebaria *et al.* (2013), who showed that English listeners have difficulty perceiving Spanish stress, especially with Spanish stress patterns that are not common in English.

Third, we cannot exclude the possibility that the differences between the Spanish and German listeners were due to the fact that the former had access to the lexical representation of the words they heard, whereas the latter did not. To test this hypothesis, we ran the above-described experiment with 14 German listeners with advanced knowledge of Spanish (mean age = 23.4 yrs, stdev = 1.5; mean age of acquisition = 14.79 yrs, stdev = 1.18; Spanish courses mean duration = 5.79 yrs, stdev = 1.93). Results showed not only that they reached the native Spanish performance (89.73% versus 90.86% for Spanish listeners), but also that the introduction of talker and intonation variability had the same impact as for the native Spanish listeners (1into-1talker = 94.44%, 1into-2talkers = 93.45%; 2into-1talker = 88.69%, 2into-2talkers = 82.34%). This finding showed that advanced knowledge of Spanish led to a native-like discrimination of Spanish stress contrasts, at least for German listeners. However, this experiment did not enable us to determine whether their native-like performance came from the existence of (Spanish) lexical representations or from their exposure to the Spanish stress patterns and their acoustic correlates.

As expected, native French listeners, in comparison with German listeners, showed severe difficulties in identifying the stress deviant, although their mean performance was still above chance level (average 44% correct responses). This result gives support for the view that the existence of lexical stress contrasts in a native language helps listeners detect such contrasts in non-native languages.

With regard to the effect of phonetic variability in the perception of lexical stress, our results showed first that the introduction of talker variability did not affect the identification of the stress deviant by Spanish listeners, which is in agreement with Dupoux *et al.* (2001). On the other hand, the introduction of intonation variability (with or without talker variability) significantly lowered the Spanish listeners' performance. This finding is in line with the results of Zahner *et al.* (2016, 2017), who found that phrase-level intonation affects the activation of stress competitors in online speech comprehension in German and in Australian English, but diverges from previous research in Spanish which suggested that listeners re-weight cues relating to lexical stress to take account of the intonation pattern (Muñoz García, 2010; Ortega-Llebaria, 2009; Ortega-Llebaria and Prieto, 2009). Two reasons might account for this discrepancy. First, we used a more cognitively demanding task in which listeners had to compare three items, while a task with lower cognitive load (i.e., identification task) was used in the previous studies. It is plausible, therefore, that the effect of intonation variability might only be observed when the task becomes more difficult. Second, the words were not presented in their original context, but were extracted and presented in isolation. It is possible that the presentation of isolated words with a rising intonation (although possible in Spanish) sounded somewhat unnatural to the listeners and thus affected their ability to identify stress.

As for the native German listeners, results showed that their detection of the stress deviant was not hampered by the introduction of talker variability, but that their performance strongly decreased with the introduction of intonation variability (independently of the presence of talker variability). It is interesting to note that, despite their lower performance in comparison with the Spanish listeners, the German listeners showed the same sensitivity to the two sources of phonetic variability (intonation and talker) as the Spanish listeners did.

The French listeners could only correctly complete the stress detection task at a level significantly above chance without the presence of intonation variability. Unlike Spanish and German listeners, the introduction of talker variability significantly decreased their performance (as in Dupoux *et al.*, 2001). In the presence of intonation variability their performance drastically dropped to chance level. We observe that phonetic variability reduces stress detection ability to a greater extent in listeners with fixed stress in their native language (here: French) than in listeners with free stress (here: German). The stronger effect of phonetic variability constitutes additional evidence in support of the view that French listeners are not able to encode the information relative to stress in another language, presumably since they have not developed this ability in their native language. This conclusion is not in agreement with Michelas *et al.* (2016), who found, in an event-related potential study, that French listeners are able to create abstract representation of stress. Here again, the discrepancy might be explained first by the different tasks, as our Odd-One-Out task is likely to be more cognitively demanding than the discrimination task in Michelas *et al.* (2016). Second, it can be explained by the knowledge that listeners had in Spanish and Italian. In our experiment, we made sure that our participants had no knowledge of another free-stress Romance language, whereas participants in Michelas *et al.* (2016) showed various levels of competence in Spanish/Italian, and hence had exposure to stress contrasts in these languages.

Despite the fact that studies in L1 have demonstrated that talker variability results in a processing cost in speech perception (e.g., Johnson, 2005; Mullennix *et al.*, 1989; Pisoni, 1997), we found that the Spanish listeners were not affected by the talker variability. An explanation might be found in Magnuson and Nusbaum's (2007) results. They showed that acoustic variability resulted in an increase in recognition times only when the variability signaled different talkers. Thus, it is possible that the talkers in our experiment sounded similar to each other. This view finds support in light of an important cue to talker identity, mean fundamental frequency, which did not significantly differ between our talkers [talker 1 = 192 Hz, talker 2 = 188 Hz; $t(22) = 1.1$, $p = 0.284$]. Another explanation that could account for the absence of talker effects in the present study—and also in Dupoux *et al.* (2001)—resides in the small number of different talkers (i.e., two), which might not have created a large processing cost for the native listeners, especially in comparison with previous studies with a higher number of talkers (e.g., 15 talkers in Mullennix *et al.*, 1989). The same explanation holds for the German listeners, since, as the Spanish listeners, they are used to stress encoding mechanisms. In contrast, the performance of the French listeners, who do not use these mechanisms in their native language, decreased with a low talker variability.

Intonation variability had a much greater influence on stress detection performance across all three listener groups, including the native group. As mentioned in Sec. I, the interrogative modality was expressed by means of rises in $F0$, which means that listeners could not rely on $F0$ to identify the position of the stressed syllable in the same way as they did for the declarative stimuli. The fact that listeners had to use this cue in a different way in interrogative stimuli seemed to hamper their perception of lexical stress, and thus the detection of the stress deviant. These findings confirm the crucial role of $F0$ in the perception of lexical stress, not only for native Spanish listeners (see Llisterri *et al.*, 2005 for similar conclusions), but also for non-native German and French listeners (see Schwab and Llisterri, 2010, 2015, for similar results for native French listeners).

The deviant detection task used in the present experiment can be related to electrophysiological research (i.e., EEG) on the perception of stress and rhythmic patterns. However, most of these studies have dealt with the identification or discrimination of stress patterns in the listeners' native language, whether this being a free- (e.g., English: Zora *et al.*, 2015; German: Domahs *et al.*, 2008) or fixed-stress language (e.g., Polish: Domahs *et al.*, 2012b; Turkish: Domahs *et al.*, 2012a; Hungarian: Honbolygó and Csépe, 2012; French: Michelas *et al.*, 2016). As for EEG studies in foreign languages, Schmidt-Kassow *et al.* (2011) showed that French advanced learners of German, contrary to German listeners, were not sensitive to rhythmic violations in German (i.e., they failed to show an electrophysiological response, P600, to metrically incorrect sentences in German). In this context, an electrophysiological study using mismatch negativity component (Näätänen *et al.*, 2007) with French and German listeners in Spanish would be particularly relevant for a better understanding of the listeners' ability to discriminate stress contrasts in a foreign language.

In conclusion, our findings not only support the view that phonetic variability has a detrimental effect on stress detection ability, but they also highlight that different sources of phonetic variability do not have the same effect on German and French listeners. On the one hand, French listeners were affected by the within-category variability induced by talkers, whereas German listeners were not. On the other hand, the intonation variability—that modified the differences between categories (i.e., stress patterns)—had an impact on stress detection ability in all listener groups, although to a greater extent in French listeners. These results indicate that some specific sources of phonetic variability (i.e., intonation) can be responsible for stress deafness in listeners who, according to the stress deafness hypothesis, are not expected to experience stress deafness in a non-native language due to the existence of contrastive stress in their native language (here: German). These findings also confirm that phonetic variability impairs stress detection ability to a greater extent in listeners with fixed stress in their native language than in listeners with free stress. All in all, the present research enables us to conclude that the processing of stress constitutes a gradient phenomenon. We observed that German listeners listening to stimuli with intonation variability (67% correct responses) revealed similar stress identification performance to French listeners listening to stimuli with no talker or intonation variability (63% correct responses). In conclusion, the present study indicates that the L1 stress characteristics not only determine the degree to which listeners are sensitive to stress, but also to what extent their stress detection abilities are impacted by distinct sources of phonetic variability.

ACKNOWLEDGMENTS

We would like to thank M. Amato, A. Bürki, L. Baqué, M[a] A. Barquero, M. Carranza, E. Dupuis, J. M. Lahoz, J. Llisterri, M. Machuca, A. Ríos, S. Schmid, and R. Tanase for their helpful advice during the elaboration and analysis of this experiment. A special thanks to F. Santiago for the ToBi transcription and to R. Rhodes for the proofreading of the manuscript. The work of S.S. was supported by the Swiss National Science Foundation (Grant No. Ambizione PZ00P1_148036/1).

[1]The underlined syllable in these examples, and in the rest of the paper, corresponds to the stressed syllable.

[2]Since Catalan is a free-stress language with similar stress patterns and intonational patterns as Spanish (Prieto, 2014), knowledge of Catalan should not have an effect on the performance of bilingual participants in Spanish. The same holds for Swiss German and German (Fitzpatrick-Cole, 1999; Fleischer and Schmid, 2006; Leemann, 2009).

[3]See Table I in supplementary material at http://dx.doi.org/10.1121/1.5008849 for the description of the trials used in the experiment.

[4]See Table II in supplementary material at http://dx.doi.org/10.1121/1.5008849 for the acoustic description of the stimuli used in the experiment.

[5]See Table III in supplementary material at http://dx.doi.org/10.1121/1.5008849 for the detailed results of the two perceptual evaluations.

[6]See Table IV supplementary material at http://dx.doi.org/10.1121/1.5008849 for the summary of a logistic mixed-effects regression model.

[7]See Table V supplementary material at http://dx.doi.org/10.1121/1.5008849 for the pairwise comparisons for the effect of phonetic variability for Spanish, German, and French listeners.

Antoniou, M., Wong, P. C. M., and Wang, S. (**2015**). "The effect of intensified language exposure on accommodating talker variability," J. Speech, Lang., Hear. Res. **58**(3), 722–727.

Baayen, R. H., Davidson, D. J., and Bates, D. M. (**2008**). "Mixed effects modeling with crossed random effects for subjects and items," J. Memory Lang. **59**, 390–412.

Baumann, S., Grice, M., and Benzmüller, R. (**2000**). "GToBI-a phonological system for the transcription of German intonation," in *Proceedings of Prosody 2000: Speech Recognition and Synthesis Workshop*, Cracow, Poland, pp. 21–28.

Delais-Roussarie, E., Post, B., Avanzi, M., Buthke, C., Di Cristo, A., Feldhausen, I., Jun, S.-A., Martin, P., Meisenburg, T., Rialland, A., Sichel-Bazin, R., and Yoo, H. (**2015**). "Intonational phonology of French:

Developing a ToBI system for French," in *Intonation in Romance*, edited by S. Frota and P. Prieto (Oxford University Press, Oxford), pp. 63–100.
Delattre, P. (**1938**). "L'accent final en français: Accent d'intensité, accent de hauteur, accent de durée (Final stress in French: Intensity, height and duration stress)," French Rev. **12**, 141–145.
Delattre, P. (**1969**). "An acoustic and articulatory study of vowel reduction in four languages," Int. Rev. Appl. Ling. Lang. Teach. **7**, 294–325.
Dogil, G., and Williams, B. (**1999**). "The phonetic manifestation of word stress," in *Word Prosodic Systems in the Languages of Europe*, edited by H. van der Hulst (Mouton de Gruyter, Berlin, New York, NY), pp. 273–334.
Domahs, U., Genc, S., Knaus, J., Wiese, R., and Kabak, B. (**2012a**). "Processing (un-)predictable word stress: ERP evidence from Turkish," Lang. Cognit. Process. **28**, 335–354.
Domahs, U., Knaus, J., Orzechowska, P., and Wiese, R. (**2012b**). "Stress 'deafness' in a language with fixed word stress: An ERP study on Polish," Front. Psychol. **3**, 1–15.
Domahs, U., Wiese, R., Bornkessel-Schlesewsky, I., and Schlesewsky, M. (**2008**). "The processing of German word stress: Evidence for the prosodic hierarchy," Phonology **25**, 1–36.
Dupoux, E., Pallier, C., Sebastián, N., and Mehler, J. (**1997**). "A destressing 'deafness' in French?," J. Memory Lang. **36**(3), 406–421.
Dupoux, E., Peperkamp, S., and Sebastián, N. (**2001**). "A robust method to study stress 'deafness'," J. Acoust. Soc. Am. **110**(3–1), 1606–1618.
Dupoux, E., Sebastián, N., Navarrete, E., and Peperkamp, S. (**2008**). "Persistent stress 'deafness': The case of French learners of Spanish," Cognition **106**(2), 682–706.
Estebas-Vilaplana, E., and Prieto, P. (**2010**). "Castilian Spanish intonation," in *Transcription of Intonation of the Spanish Language*, edited by P. Prieto and P. Roseano (Lincom Europa, Munich, Germany), pp. 17–48.
Fitzpatrick-Cole, J. (**1999**). "The alpine intonation of Bern Swiss German," in *Proceedings of the XIVth International Congress of Phonetic Sciences* (*ICPhS*), San Francisco, CA (August 1–7).
Fleischer, J., and Schmid, S. (**2006**). "Zurich German," J. Int. Phonetic Assoc. **36**(2), 243–255.
Hayes, B. (**1995**). *Metrical Stress Theory: Principles and Case Studies* (University of Chicago Press, Chicago, IL), pp. 1–473.
Hirst, D. (**2011**). "The analysis by synthesis of speech melody: From data to models," J. Speech Sciences **1**(1), 55–83.
Honbolygó, F., and Csépe, V. (**2012**). "Saliency or template? ERP evidence for long-term representation of word stress," Int. J. Psychophysiol. **87**(2), 165–172.
Hualde, J. I., and Prieto, P. (**2015**). "Intonational variation in Spanish: European and American varieties," in *Intonational Variation in Romance*, edited by S. Frota and P. Prieto (Oxford University Press, Oxford), pp. 350–391.
Hyman, L. M. (**2006**). "Word-prosodic typology," Phonology **23**, 225–257.
Janssen, U. (**2003**). "Untersuchungen zum Wortakzent im Deutschen und Niederländischen (Research on lexical stress in German and Dutch)," Doctoral thesis, Universität Düsseldorf.
Jessen, M., Marasek, K., Schneider, K., and Clahßen, K. (**1995**). "Acoustic correlates of word stress and the tense/lax opposition in the vowel system of German," Proc. Int. Congress Phonetic Sci. **13**, 428–431.
Johnson, K. (**2005**). "Speaker normalization in speech perception," in *The Handbook of Speech Perception*, edited by D. B. Pisoni and R. Remez (Blackwell Publishers, Oxford), pp. 363–389.
Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (**2014**). "lmerTest: Tests in linear mixed effects models. R package version 2.0-20," URL http://CRAN.R-project.org/package=lmerTest (Last viewed October 24, 2016).
Lacheret-Dujour, A., and Beaugendre, F. (**1999**). *La Prosodie du Français (French Prosody)* (CNRS éditions, Paris, France), pp. 1–354.
Laganaro, M., Vacheresse, F., and Frauenfelder, U. H. (**2002**). "Selective impairment of lexical stress assignment in an Italian-speaking aphasic patient," Brain Lang. **81**, 601–609.
Leemann, A. (**2009**). "Intonational variation in Swiss German," in *Language Variation. European Perspectives*, edited by F. Hinskens (Benjamins, Amsterdam/Philadelphia), pp. 135–143.
Léon, P. (**2007**). *Phonétisme et Prononciations du Français* (French Phonetics and Pronunciations) (Armand Colin, Paris), pp. 1–269.
Levelt, W. J. M., Roelofs, A., and Meyer, A. S. (**1999**). "A theory of lexical access in speech production," Behav. Brain Sci. **22**, 1–38.
Llisterri, J., Machuca, M. J., Mota, C., de la Riera, M., and Ríos, A. (**2005**). "La percepción del acento léxico en español (The perception of lexical stress in Spanish)," in *Filología y Lingüística. Estudios Ofrecidos a Antonio Quilis* (Consejo Superior de Investigaciones Científicas—Universidad Nacional de Educación a Distancia—Universidad de Valladolid, Madrid, Spain), pp. 271–297.
Llisterri, J., Machuca, M. J., Ríos, A., and Schwab, S. (**2014**). "El acento en contexto: Datos acústicos (Stress in context: Acoustic data)," in *Fonética Experimental, Educación Superior e Investigación* (*Experimental Phonetics, Superior Education and Investigation*), edited by Y. Congosto Martín, M. L. Montero Curiel, and A. Salvador Plans (Arco/Libros, Madrid), Vol. I, pp. 357–376.
Magnuson, J. S., and Nusbaum, H. C. (**2007**). "Acoustic differences, listener expectations, and the perceptual accommodation of talker variability," J. Exp. Psychol. **33**, 391–409.
Michelas, A., Frauenfelder, U. A., Schön, D., and Dufour, S. (**2016**). "How deaf are French speakers to stress?," J. Acoust. Soc. Am. **139**, 1333–1342.
Mooshammer, C., and Geng, C. (**2008**). "Acoustic and articulatory manifestations of vowel reduction in German," J. Int. Phon. Assoc. **38**, 117–136.
Mullennix, J. W., Pisoni, D. B., and Martin, C. S. (**1989**). "Some effects of talker variability on spoken word recognition," J. Acoust. Soc. Am. **85**, 365–378.
Muñoz García, M. (**2010**). "La perception et la production de l'accent lexical de l'espagnol par des francophones: Aspects phonétiques et psycholinguistiques (The perception and production of Spanish lexical stress by French speakers: Phonetic and psycholinguistic aspects)," Doctoral thesis, Université de Toulouse.
Näätänen, R., Paavilainen, P., Rinne, T., and Alho, K. (**2007**). "The mismatch negativity (MMN) in basic research of central auditory processing: A review," Clin. Neurophysiol. **118**, 2544–2590.
Nieuwenhuis, R., Pelzer, B., and te Grotenhuis, M. (**2012**). "Tools for detecting influential data in mixed effects models," R J. **4**(2), 38–47. URL https://journal.r-project.org/archive/2012-2/RJournal_2012-2_Nieuwenhuis~et~al.pdf (Last viewed October 24, 2016).
Ortega-Llebaria, M. (**2009**). "On-line perception of lexical stress in English and Spanish: The effects of sentence intonation and vowel reduction," J. Acoust. Soc. Am. **125**(4), 2755.
Ortega-Llebaria, M., Gu, H., and Fan, J. (**2013**). "English speakers' perception of Spanish lexical stress: Context-driven L2 stress perception," J. Phon. **41**, 186–197.
Ortega-Llebaria, M., and Prieto, P. (**2009**). "Perception of word stress in Castilian Spanish: The effects of sentence intonation and vowel type," in *Phonetics and Phonology. Interactions and Interrelations*, edited by M. Vigário, S. Frota, and M. J. Freitas (John Benjamins Publishing Company, Amsterdam), pp. 35–50.
Peperkamp, S., Vendelin, I., and Dupoux, E. (**2010**). "Perception of predictable stress: A cross-linguistic investigation," J. Phonetics **38**(3), 422–430.
Pisoni D. B. (**1997**). "Some thoughts on 'normalization in speech perception'," in *Talker Variability in Speech Processing*, edited by K. Johnson and J. W. Mullennix (Academic Press, San Diego, CA), pp. 9–32.
Prieto, P. (**2014**). "The intonational phonology of Catalan," in *Prosodic Typology II: The Phonology of Intonation and Phrasing*, edited by S.-A. Jun (OUP, Oxford), pp. 43–80.
Protopapas, A., Panagaki, E., Andrikopoulou, A., Gutiérrez Palma, N., and Arvaniti, A. (**2016**). "Priming stress patterns in word recognition," J. Exp. Psychol. **42**(11), 1739–1760.
Quilis, A. (**1981**). *Fonética Acústica de la Lengua Española (Acoustic Phonetics in Spanish)* (Gredos, Madrid), pp. 1–502.
R Development Core Team (**2016**). "R: A language and environment for statistical computing," *R Foundation for Statistical Computing*, Vienna, Austria, URL http://www.R-project.org/.
Rahmani, H., Rietveld, T., and Gussenhoven, C. (**2015**). "Stress 'deafness' reveals absence of lexical marking of stress or tone in the adult grammar," PLoS One **10**(12), e0143968.
Rossi, M. (**1979**). "Le français, langue sans accent (French, language without stress)," Studia Phonetica **15**, 13–52.
Santiago Vargas, F., and Delais-Roussarie, E. (**2012**). "Acquiring phrasing and intonation in French as second language: The case of yes-no questions produced by Mexican Spanish learners," in *Proceedings of the 6th International Conference on Speech Prosody*, Shanghai, China (May 22–25).
Schmidt-Kassow, M., Rothermich, K., Schwartze, M., and Kotz, S. A. (**2011**). "Did you get the beat? Late proficient French-German learners extract strong–weak patterns in tonal but not in linguistic sequences," NeuroImage **54**, 568–576.

Schwab, S., and Dellwo, V. (**2016**). “The use of the Odd-One-Out task in the study of the perception of lexical stress in Spanish by native German listeners?,” in *Proceedings of the 8th Speech Prosody*, Boston, MA (May 31–June 3).

Schwab, S., and Llisterri, J. (**2010**). “La perception de l’accent lexical espagnol par des apprenants francophones (The perception of Spanish lexical stress by French learners),” in *La langue et l’être communiquant. Hommage à Julio Murillo (Language and Communicating being. In honor to Julio Murillo)*, edited by L. Baqué and M. Estrada (Éditions du CIPA, Mons), pp. 311–328.

Schwab, S., and Llisterri, J. (**2015**). “The discrimination of Spanish lexical stress contrasts by French-speaking listeners,” *Proceedings of the International Symposium of Monolingual and Bilingual Speech*, Chania, Greece (September 7–10).

Tremblay, A. (**2009**). “Phonetic variability and the variable perception of L2 word stress by French Canadian listeners,” Int. J. Bilingualism **13**, 35–62.

Vaissière, J. (**2002**). “Cross-linguistic prosodic transcription: French vs English,” in *Problems and Methods of Experimental Phonetics. In Honour of the 70th Anniversary of Pr. L. V. Bondarko*, edited by N. B. Volskaya, N. D. Svetozarova, and P. A. Skrelin (St. Petersburg State University Press, St. Petersburg), pp. 147–164.

van Dommelen, W. A., and Hazan, V. (**2012**). “Impact of talker variability on word recognition in non-native listeners,” J. Acoust. Soc. Am. **132**, 1690–1699.

Wiese, R. (**1996**). *The Phonology of German* (Clarendon Press, Oxford), pp. 1–359.

Wochner, D., Schlegel, J., Braun, B., and Dehé, N. (**2015**). “The prosody of rhetorical questions in German,” in *Proceedings of Interspeech*, Dresden, Germany (September 6–10).

Zahner, K., Kember, H., and Braun, B. (**2017**). “Mind the peak: When museum is temporarily understood as musical in Australian English,” in *Proceedings of Interspeech*, Stockholm, Sweden (August 20–24).

Zahner, K., Schönhuber, M., Grijzenhout, J., and Braun, B. (**2016**). “Pitch accent type affects lexical activation in German: Evidence from eye tracking,” in *Proceedings of the 16th Speech Science and Technology Conference* (SST 2016), Sydney, Australia (December 7–9, 2016).

Zora, H., Schwarz, I.-C., and Heldner, M. (**2015**). “Neural correlates of lexical stress: Mismatch negativity reflects fundamental frequency and intensity,” Neuroreport **26**(13), 791–796.