L2 stress discrimination by non-musicians and musicians playing wind or percussion instruments

Sandra SCHWAB
https://orcid.org/ 0000-0003-4485-8335; sandra.schwab@unibe.ch
University of Bern (SWITZERLAND)

Orianne PYTHOUD
orianne.pythoud@unifr.ch
University of Fribourg (SWITZERLAND)

Suggested citation: SCHWAB, S., PYTHOUD, O. (2023), L2 stress discrimination by non-musicians and musicians playing wind or percussion instruments, *Langue(s) & Parole,* 8, 179-193 https://doi.org/10.5565/rev/languesparole.133

Abstract

The present study explores the influence of music expertise on French-speaking listeners' ability to process L2 Spanish stress. Musicians playing wind or percussion instruments and non-musicians completed an Odd-One-Out task, where they heard three Spanish words presented in babble noise and were asked to detect the word with a different stress pattern. Results first showed that musicians (playing either wind or percussion instruments) outperformed non-musicians, confirming the advantage of music expertise in 'speech in noise' perception. Secondly, they revealed that percussionists – who, as rhythm experts, relied more on stress-related timing cues – performed better than musicians playing wind instruments – who, as pitch experts, were more sensitive to stress-related pitch cues – in detecting stress in Spanish words presented in babble noise. Finally, there was no evidence of the larger advantage of being a musician when processing L2 stress in trials produced with a large degree of phonetic variability. Taken together, our findings not only highlighted the benefits of music expertise in L2 stress perception, but also revealed that the type of music instruments played by the musicians also influences L2 stress discrimination performance.

Keywords: Music expertise, word stress perception, second language, babble noise

Résumé

Cette étude explore l'influence de l'expertise musicale sur la capacité de participants francophones à percevoir l'accent lexical en espagnol L2. Des musiciens jouant des instruments à vent ou des percussions et des non-musiciens ont effectué à une tâche de "Odd-One-Out" dans laquelle ils entendaient trois mots espagnols présentés dans du bruit de fond et devaient détecter le mot avec un patron accentuel différent. Les résultats ont premièrement montré que les musiciens (qu'ils jouent des instruments à vent ou des percussions) étaient meilleurs que les non-musiciens, confirmant ainsi l'avantage de l'expertise musicale dans la perception de la parole dans le bruit. Deuxièmement, ils ont révélé que les percussionnistes - experts en rythme, s'appuyant davantage sur les indices temporels liés à l'accent - obtenaient de meilleurs résultats que les musiciens jouant des instruments à vent - experts en variations mélodiques - pour détecter l'accent dans les mots espagnols présentés dans du bruit de fond. Enfin, les résultats n'ont pas montré que l'expertise musicale constituait un avantage plus important lorsque les auditeurs devaient discriminer des mots produits avec un degré élevé de variabilité phonétique. Dans l'ensemble, nos résultats ont donc mis en évidence les avantages de l'expertise musicale dans la perception de l'accent lexical en L2 et ont révélé que le type d'instrument de musique joué par les musiciens influence aussi les capacités de ces derniers à discriminer les contrastes accentuels en L2.

Mots-clés : Expertise musicale, perception de l'accent lexical, langue seconde, bruit de fond

Resumen

La presente investigación explora la influencia de la experiencia musical en la capacidad de participantes francófonos para percibir el acento léxico en español como segunda lengua (L2). Músicos que tocaban instrumentos de viento o percusión y no músicos participaron en una tarea de "Odd-One-Out" en la que escuchaban tres palabras españolas presentadas con ruido de fondo y debían detectar la palabra con un patrón acentual diferente. Los resultados mostraron, en primer lugar, que los músicos (ya sea que tocaran instrumentos de viento o de percusión) eran mejores que los no músicos, confirmando así la ventaja de la experiencia musical en la percepción del habla en el ruido. En segundo lugar, revelaron que los percusionistas –expertos en ritmo que se apoyaban más en los indicios temporales relacionados con el acento— obtuvieron mejores resultados que los músicos que tocan instrumentos de viento -expertos en variaciones melódicas— en detectar el acento en las palabras españolas presentadas con ruido de fondo. Finalmente, los resultados no mostraron que la experiencia musical proporcionara una ventaja más significativa cuando los oyentes tenían que discriminar palabras producidas con un alto grado de variabilidad fonética. En conjunto, nuestros resultados destacan las ventajas de la experiencia musical en la percepción del acento léxico en L2 y revelan que el tipo de instrumento de música tocado por los músicos también influye en la discriminación de contrastes acentuales en L2.

Palabras clave: Experiencia musical, percepción del acento léxico, segunda lengua, ruido de fondo

Resum

La present investigació explora la influència de l'experiència musical en la capacitat dels participants francòfons per percebre l'accent lèxic en espanyol com a segona llengua (L2). Músics que tocaven instruments de vent o percussió i no músics van participar en una tasca de "Odd-One-Out" en què escoltaven tres paraules espanyoles presentades amb soroll de fons i havien de detectar la paraula amb un patró accentual diferent. Els resultats van mostrar, en primer lloc, que els músics (tant si tocaven instruments de vent com de percussió) eren millors que els no músics, confirmant així l'avantatge de l'experiència musical en la percepció de la parla en soroll. En segon lloc, van revelar que els percussionistes –experts en ritme que es recolzaven més en els indicis temporals relacionats amb l'accent— van obtenir millors resultats que els músics que tocaven instruments de vent –experts en variacions melòdiques— en detectar l'accent en les paraules espanyoles presentades amb soroll de fons. Finalment, els resultats no van mostrar que l'experiència musical proporcionés un avantatge més significatiu quan els oients havien de discriminar paraules produïdes amb un alt grau de variabilitat fonètica. En conjunt, els nostres resultats destaquen els avantatges de l'experiència musical en la percepció de l'accent lèxic en L2 i revelen que el tipus d'instrument de música tocat pels músics també influeix en la discriminació de contrastos accentuals en L2.

Paraules clau: Experiència musical, percepció de l'accent lèxic, segona llengua, soroll de fons.

1. Introduction

1.1 L2 stress perception and music expertise

Prosody can be defined as 'the music of speech', which could be simplistically considered as a combination of rhythm and melody. The present research deals with one specific prosodic feature, accentuation, which is essential for creating rhythmic patterns in speech. More specifically, focus will be on the perception of word stress in a second/foreign language (L2).

Word stress is the accentuation of a syllable within a word. Perceptually, the stressed syllable is more salient than the unstressed syllables of a word. For example, the word 'history' has stress on the first syllable, whereas 'historical' has stress on the second syllable. Although language-specific, the perceptual correlates of word stress involve variations in pitch, intensity and duration (e.g., Llisterri *et al.*, 2003 about perceptual correlates of Spanish stress). Hence, importantly, word stress can be perceived based on rhythmic and/or melodic modulations in speech.

Besides tonal languages, languages can be classified into 'free-stress' languages and 'fixed-stress' languages. In the former category (e.g., Spanish, English, German), the position of word stress is defined by morphophonological constraints and is said to be 'variable'. In English, for example, some words bear stress on the first syllable (e.g., prosody)[1], others on the second syllable (e.g., prosodic) or on the final syllable (e.g., engineer). Word stress in these languages is distinctive, since it distinguishes the meaning of two words (e.g., es. número vs. numero, en. (the) number vs I number). In fixed-stress languages (e.g., Hungarian, French), the position of stress is fixed and generally predictable. For example, Hungarian word stress in disyllabic words always falls on the first syllable (Honbolygó *et al.* 2020). Hence, word stress in fixed-stress languages does not fulfill a distinctive function.

The perception of word stress in a foreign language with free stress (e.g., English, Spanish) has been shown to be complicated for listeners whose native language bears fixed stress (e.g., French, Hungarian). This perceptual difficulty is explained by the listeners' inability to process word stress phonologically in the foreign language (i.e., as a distinctive feature), since word stress is not distinctive in their native language (Dupoux *et al.*, 1997; Schwab, Dellwo, 2017).

It is interesting to underline that the parameters employed in the perception of word stress - pitch, duration, intensity - are also used in music. Given these similarities, it seems legitimate to ask whether music expertise would help listeners, especially with fixed stress languages (e.g., French), to process word stress in a foreign language with free stress (e.g., Spanish, English).

A couple of studies have indeed compared L2 stress processing by French-speaking non-musicians and musicians. Schwab, Calpini, 2018, for example, showed that musicians (with an average of 30 years of music training) outperformed non-musicians in identifying the stressed syllable in L2 Spanish words. Kolinsky *et al.* (2009) tested French-speaking musicians (with at least 10 years of music training) and non-musicians in a sequence repetition task using L2 English pseudowords presenting weak or strong word stress contrasts. The musicians showed higher performance

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

compared to the non-musicians, particularly in the condition with weak stress contrasts. These findings suggest that music expertise may be especially advantageous in situations where the detection of word stress is perceptually difficult.

Speech perception in noise is a challenging task that requires separating the target signal from competing background noise. Musicians, thanks to their extensive training, are highly capable of extracting relevant signals from complex soundscapes (e.g., distinguishing their own instrument's sound in an orchestra). The effect of such capacity has been shown to be positively transferred to the listeners' abilities to perceive speech in noise. It has been demonstrated that musicians clearly have an advantage over non-musicians in 'speech in noise' perception, at least regarding sentence comprehension (e.g., Baçent, Gaudrain, 2016; Kumar, Krishna, 2019; Swaminathan *et al.*, 2015).

Nevertheless, from all of the aforementioned studies, one question still has to be answered. Since no distinctions were made among the categories of professional music performers (e.g., conductors, orchestra members, soloists, composers, music teachers) or the types of instruments played by the musicians (e.g., string, wind, percussion), little is known about the potentially differential influence of the distinct music activities on L2 stress perception. Yet, there is some evidence that musicians perceive pitch and timing variations differently according to the type of music instruments they play. For example, Micheyl *et al.* (2006) argued that musicians who self-tune their instruments (e.g., string, wind) were more sensitive to smaller pitch differences than musicians who did not (e.g., percussion, keyboard). Along the same lines, percussionists, considered as rhythm-experts, have been shown to be more sensitive to smaller timing deviations compared to other musicians (Ehrlé, Samson, 2005). It is thus plausible to argue that the listeners' ability to process word stress in a foreign language depends on the practice of some specific music instruments.

2. The present study

The present study investigated the influence of the listeners' music expertise on their ability to process stress in a free-stress foreign language. Particularly, it focused on word stress perception in Spanish as an L2 by

French-speaking listeners with no knowledge of Spanish (i.e., *ab initio* learners of Spanish). Participants completed an Odd-One-Out task, where they heard three Spanish words and were asked to detect the word with a different stress pattern (i.e., 'odd' word; Schwab, Dellwo, 2017).

The first aim of the present study was to examine whether musicians' advantage over non-musicians in processing speech in noise was also found during L2 stress perception. For this, the stimuli were presented in babble noise. According to the aforementioned previous research, we expected musicians to perform better than non-musicians, because the former were supposedly more able to extract the relevant linguistic information from the background noise than the latter.

In addition to elucidating the effect of music expertise on stress processing in noise, we examined whether this effect varied as a function of the type of music instruments played by the musicians. Based on the findings that musicians who play music instruments that are rather pitch-focused (e.g., string, wind) or rhythm-focused (e.g., percussion) have different pitch and timing perception, the second goal of the present research was to determine whether musicians playing wind instruments and percussionists differed in their ability to process word stress in a foreign language. As mentioned earlier, word stress can be perceptually characterized by a combination of pitch and timing variation. Thus, if L2 stress processing is rather pitch-based, we expected wind instrument musicians - relying more on pitch differences – to perform better than the percussionists. On the other hand, if L2 stress processing is rather rhythm-based, we predicted the percussionists – experts in timing variations – to outperform the wind instrument musicians.

Considering that music expertise has a positive impact on L2 lexical perception, especially under particularly challenging perceptual conditions (Kolinsky *et al.* 2009), our third goal was to investigate the effect of phonetic variability - making word stress processing more challenging - in the three groups under study (i.e., non-musicians, wind and percussion instrument musicians). For this, we compared the listeners' performance in trials where the three Spanish words were produced with either the same voice or with two different voices and carried the same or varying intonation contours (like in Schwab, Dellwo, 2017). We predicted the differences between musicians (wind and percussion) and non-musicians to be larger in trials with more phonetic variability, since music expertise was expected to enhance L2 stress perception in more challenging situations.

2. Methods

2.1 Participants

Forty-five participants, divided into three groups, took part in the experiment. All participants were native speakers of French, with no knowledge Spanish or of another free-stress Romance language (e.g., Italian, Portuguese). The first group was composed of 15 non-musicians (9 females; mean age = 25.13 years, stdv = 1.88). Non-musicians had no musical training, except the one gathered from their time in compulsory school (1 hour per week, mainly singing practice), they did not play any instrument nor did they sing in a choir. The second group comprised 15 wind instrumentalists (9 females; mean age = 24.13 years, stdv = 1.34) including 7 people playing instruments from the brass section (trombone, trumpet, tenor horn) and 8 from the woodwind category (clarinet, flute, saxophone, bassoon, oboe). They begun to play the instrument at the average age of 8.27 years (stdv = 2.02) and had an average musical training of 13.8 years (stdv = 2.14). All had ensemble experience (e.g., orchestra, brass band). The third group consisted of 15 percussionists (7 females; mean age = 26.73 years, stdv = 4.23) who all played a whole range of percussion instruments (e.g., drums, timpani, claves, marimba, xylophone). They had begun to play the instruments at the average age of 8.87 years (stdv = 2.67) and had an average musical training of 13.13 years (stdv = 2.56). All of them were also part of an ensemble.

2.2 Material

2.2.1 Stimuli

Participants heard trials of three trisyllabic Spanish words and had to indicate the word with a different stress pattern (i.e., 'odd' word). The material and procedure were taken from Schwab, Dellwo (2017). Six trisyllabic Spanish words were used in the experiment. These were 'numero' and 'valido', which were produced in their proparoxytone form (i.e., first syllable stressed; número and válido), paroxytone (i.e., second syllable stressed: numero and valido) and oxytone (final syllable stressed: numeró, validó). The experiment was composed of 216 trials composed of three words separated by 500 ms. Two of the words had the same stress pattern and one a different stress pattern (i.e., the 'odd'). For example, in the trial 'número-número-numero', the deviant ('odd') word was the third word (with stress on the second syllable, while the two first words had stress on the first syllable).

Different degrees of phonetic variability, introduced by voice and/or intonation variability, were present in the trials. Trials comprised words produced by one voice, with only falling intonation ('1voice1into'), while other trials were composed of words produced by two voices, with again only falling intonation ('2voices1into'). Similarly, trials also contained words produced by one voice, but this time pronounced with falling and rising intonations ('1voice2into'), and other trials consisted of words produced by two voices, again pronounced with falling and rising intonations. ('2voices2into'). As for the acoustic cues used to signal word stress, Schwab, Dellwo (2017) reported that word stress was realized by variation in both fundamental frequency (i.e., acoustic correlate of pitch) and duration, with the former being less relevant than the latter in words with rising intonation.

2.2.2 Babble noise

A babble noise was superimposed on each trial. The babble noise was taken from a Youtube video entitled "Ambiance soirée discussion sans musique" (https://www.youtube.com/watch?v=GS3hkkg-pqM&t=15s) and contained several overlapping voices speaking in French (i.e., the participants' native language). The babble noise began 100 ms before the beginning of each trial and was played continuously during the entire duration of the trial (i.e., also during the pauses between the words). Following Başkent, Gaudrain (2016), the signal-to-noise ratio (i.e., ratio between the volume in decibels of the trial and that of the babble noise) was set at -6 dB. With trials normalized at 66 dB, the babble noise was set at 60 dB.

2.3 Procedure

The experiment was administered with Praat software (Boersma, Weenink 2011). Participants performed the experiment individually in a quiet environment, on a Macintosh laptop with Marshall headphones (Major IV model). Before beginning the experiment, the participants were informed about the presence of a babble noise superimposed on the trials and completed practice trials in order not to be surprised by the babble noise and to get familiar with the task. After hearing each trial (with the babble noise), the participants were asked to indicate which of the three words was the deviant one, i.e., which word had a different stress pattern, by clicking on the corresponding response on the screen. The experiment lasted approximately 25 minutes.

2.4 Data analysis

Data of four participants (3 NonMusician, 1 Wind, 1 Percussion) were removed due to technical issues. The statistical analysis was conducted using R software (version 4.0.3; R Development Core Team, 2022) with the lme4 R package (Bates, Mächler, Bolker, Walker, 2015). A binary response mixed-effects logistic regression was used to model the correct/incorrect responses at the Odd-One-Out task (Baayen, Davidson, Bates, 2008). The fixed part of the model was comprised of 'group' (NonMusician, Wind, Percussion) and 'phonetic variability' (1voice1into, 2voices1into, 2voices1into, 2voices2into) and the two-way interaction between both variables. The random part of the model included random intercepts for participants and items. The random slope allowing for the effect of 'phonetic variability' to differ across participants was not included due to singularity issues. Likelihood ratio tests were used to determine the significance of the main and interaction effects by comparing models with and without these effects. The estimates (β) were calculated in logit, with the reference level for the dependent variable being 'incorrect response'. Within Cook's space, four potential influential values were observed. These values were kept given that their exclusion did not change the results of the model. No issue related to overdispersion was observed. The figures presented in the next section display the percentage of correct responses, while all statistical analyses were conducted on the raw data (correct/incorrect responses).

3. Results and discussion

Figure 1 presents the percent correct as a function of the group (NonMusician, Wind, Percussion). A group effect is observed (see Table 1 in Appendix). Post-hoc analyses with Tukey correction ($p < .05$) showed that the performance of the three groups differed from each other, with the musicians playing percussions performing better (69.24%) than the musicians playing wind instruments (54.11%), who in turn were better than the non-musicians (36.75%).

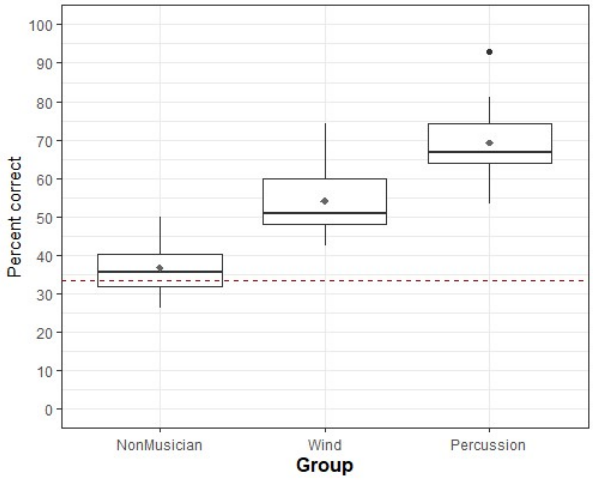

Figure 1: Percent correct as a function of group (NonMusician, Wind, Percussion). Gray diamonds represent the mean percent correct of each group and the dashed line chance level (33%).

Figure 2 presents the percent correct as a function of the group (NonMusician, Wind, Percussion) and phonetic variability (1voice1into, 2voices1into, 1voice2into, 2voices2into). We observe an effect of phonetic variability (see Table 1 in Appendix). Post-hoc comparisons with Tukey correction ($p < .05$) revealed that (for all groups) for trials composed of only one intonation pattern, the performance was similar in trials produced by one voice (1voice1into) or two voices (2voices1into). Similarly, for trials with two intonation patterns, the performance did not differ in trials produced by one voice (1voice2into) or two voices (2voices2into). In contrast, we observe that the performance significantly dropped in trials comprising two intonation patterns compared to trials produced with only one intonation pattern. These findings indicate that, contrary to intonation variability, voice variability within the trials did not affect the detection of the deviant word. Moreover, as depicted in Figure 2, the presence of the interaction (see Table 1 in Appendix) reveals that, contrary to our predictions, the differences between the groups, although still significant, decrease with intonation variability (especially in 2voices2into).

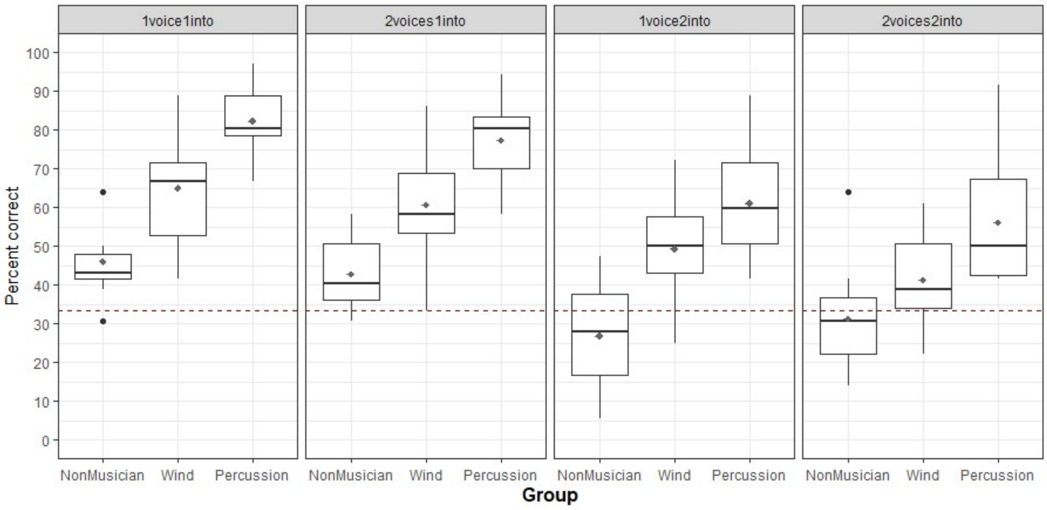

Figure 2: Percent correct as a function of group (NonMusician, Wind, Percussion) and phonetic variability (1voice1into, 2voices1into, 1voice2into, 2voices2into). Gray diamonds represent the mean percent correct for each group and phonetic variability and the dashed line chance level (33%).

4. General discussion

The present study investigated the influence of music expertise on French-speaking listeners' ability to process Spanish stress in words presented in babble noise. Our results first showed that, as expected, musicians (playing either wind or percussion instruments) performed better than non-musicians. The non-musicians' lower performance (36.75%) compared to previous studies (44.62% in Schwab,Dellwo, 2017) can be explained by the presence of the babble noise, which made the task particularly challenging for non-musicians. Overall, our findings confirmed the advantage of music expertise in 'speech in noise' perception and expanded the results found for sentence comprehension to prosodic processing (in L2).

Secondly, our findings revealed that the percussionists outperformed the musicians playing wind instruments. Considering that the former were rather rhythm-experts, while the latter more pitch-experts, and that word stress was realized with duration (i.e., timing) and f0 (i.e., pitch variations), one could be tempted to conclude that L2 stress processing was performed rather timing-based than pitch-based. Yet, this conclusion should be softened, because the percussionists' outperformance might be caused by the specific characteristics of the stimuli used in the experiment. It is indeed possible that the babble noise superimposed on the stimuli masked in a greater extent the stress-related pitch variations – that the wind instrument musicians are said to be more sensitive to – than the duration variations – that percussionists seem to rely more on – and, as a consequence, provided advantage to the percussionists. In other words, it might be that, because of the babble noise, timing cues to signal word stress were more salient than pitch cues, enhancing thus rhythm-based stress processing and benefitting percussionists. To confirm this hypothesis, similar experiments should be run without babble noise, with, however some ceiling effect risk, especially in musicians.

Third, regarding the effect of phonetic variability, our results showed that, contrary to Schwab,Dellwo (2027)' finding, voice variability did not affect stress discrimination performance. This discrepancy can be explained by the fact that our listeners have probably perceived all words as produced by only one voice, because of the presence of the babble noise that diminished the differences between the two voices. In contrast, an effect of intonation variability was observed independently of music expertise, with a general lower performance when trials were composed of words with different intonation contours. It appears that the rising intonation contour, together with the babble noise, has considerably masked the stress-related pitch cues to the point of hampering all listeners' ability to discriminate words with rising and falling intonation contours. However, contrary to our prediction, there was no evidence that music expertise had a larger advantage in processing word stress in stimuli with more phonetic variability. Although percussionists outperformed, independently of the phonetic variability, the musicians playing wind instruments, who in turn performed better than the non-musicians, the differences between the three groups (although still significant) tended to decrease with increasing phonetic variability. Thus, in light of our results, it appears that music expertise did not help more the listeners when they had to process trials produced with a large degree of phonetic variability. The disagreement with Kolinsky *et al.* (2009)' results might come from methodological differences, especially from the use, in our experiment, of babble noise.

In conclusion, the present research not only highlighted the advantage of music expertise for L2 stress perception in noise, but it also revealed that the L2 stress discrimination abilities varied depending on the type of music instruments played by the musicians. Percussionists (i.e., rhythm experts) performed better than musicians playing wind instruments (i.e., pitch experts), suggesting that, given our specific experimental design, relying rather on stress-related timing than stress-related pitch cues was more efficient for detecting stress in Spanish words presented in babble noise.

References

BAAYEN, R. H., DAVIDSON, D. J., BATES, D. M., Mixed effects modeling with crossed random effects for subjects and items, *J. MEMORY LANG*, 2008, **59**, 390-412. https://doi.org/10.1016 /j.jml.2007.12.005

BAÇENT, D., GAUDRAIN, E., Musician advantage for speech-on-speech perception, *THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA*, 2016, **139**, 51-56. https://doi.org/10.1121/1.4942628

BATES, D., MÄCHLER, M., BOLKER, B., WALKER, S., Fitting linear mixed-effects models using lme4, *JOURNAL OF STATISTICAL SOFTWARE*, 2015, **67(1)**, 1-48. doi: 10.18637/jss.v067.i01.

BOERSMA, P., WEENINK, D., *Praat: Doing Phonetics by Computer* [Computer Software], 2011. http://www.praat.org.

DUPOUX, E., PALLIER, C., SEBASTIÁN-GALLÉS, N., MEHLER, J., A destressing 'deafness' in French? *JOURNAL OF MEMORY AND LANGUAGE*, 1997, **36(3)**, 406–421. https://doi.org/10.1006/jmla.1996.2500

EHRLÉ, N., SAMSON, S., Auditory discrimination of anisochrony: Influence of the tempo and musical backgrounds of listeners, *BRAIN AND COGNITION*, 2005, **58(1)**, 133–147. https://doi.org/ 10.1016/ j.bandc.2004.09.014

HONBOLYGÓ, F., KÓBOR, A., HERMANN, P, KETTINGER, A. O, VIDNYÁNSZKY, Z., KOVÁCS, G., CSÉPE, V., Expectations about word stress modulate neural activity in speech-sensitive cortical areas, *NEUROPSYCHOLOGIA*, 2020, **143**. 107467, ISSN 0028-3932, https://doi.org/10.1016/j.neuropsychologia.2020.107467.

KOLINSKY, R., CUVELIER, H., GOETRY, V., PERETZ, I., MORAIS, J., Music training facilitates lexical stress processing, *MUSIC PERCEPTION*, 2009, **26(3)**, 235-246. https://doi.org/10.1525/mp.2009.26.3.235

KUMAR, P.V., KRISHNA, R., Exploring music induced auditory processing differences among vocalists, violinists and non-musicians, *INTERNATIONAL JOURNAL OF HEALTH SCIENCES AND RESEARCH*, 2019, **9(2)**, 13-21.

LLISTERRI, J., MACHUCA, M., DE LA MOTA, C., RIERA, M., RÍOS, A., The perception of lexical stress in Spanish, *Proceedings of the XVth International Conference of Phonetic Sciences*, 2003 ISBN. 2023-2026.

MICHEYL, C., DELHOMMEAU, K., PERROT, X., OXENHAM, A., Influence of musical and psychoacoustical training on pitch discrimination, *HEARING RESEARCH*, 2006, **219**, 36-47. https://doi.org/10.1016/j.heares.2006.05.004.

R CORE TEAM, R, *A language and environment for statistical computing. R Foundation for Statistical Computing,* Vienna, Austria, 2022. URL https://www.R-project.org/.

SCHWAB, S., CALPINI, N., Expertise musicale et perception de variations de f0 en L1 et en L2, *REVUE FRANÇAISE DE LINGUISTIQUE APPLIQUÉE*, 2018, **23(1)**, 15-30. https://doi.org/10.3917/rfla.231.0015

SCHWAB, S., DELLWO, V., Intonation and talker variability in the discrimination of Spanish lexical stress contrasts by Spanish, German and French listeners, *THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA,* 2017, **142**, 2419-2429. https://doi.org/10.1121/1.5008849

SWAMINATHAN, J., MASON, C., STREETER, T., BEST, V., KIDD, G., PATEL, A. D., Musical training, individual differences and the cocktail party problem, *SCIENTIFIC REPORT*, 2015, **5**, 1-10. https://doi.org/10.1038/srep11628

Sandra SCHWAB is a senior researcher and lecturer in French Linguistics at the University of Bern and works in parallel as a scientific collaborator at Linguistic Research Infrastructure (LiRI) of the University of Zurich. Her recent research mainly deals with the study of behavioral factors and neural correlates involved in L2 word stress perception.

Orianne PYTHOUD earned her master's degree in history and French at the University of Fribourg in June 2023. Her practice of music (flute, amateur certificate at the Fribourg Conservatory) led her to take an interest in the contribution of musical skills to speech, particularly in the context of the acquisition of a second language.

APPENDIX

Predictors	*Log-Odds*	*CI*	*p*	*Likelihood ratio tests*
(Intercept)	-0.19	-0.60 – 0.21	0.347	
Group				$\chi^2(2) = 46.11$, p < .001
Percussion	1.99	1.51 – 2.47	<0.001	
Wind	0.91	0.45 – 1.37	<0.001	
Phonetic variability				$\chi^2(3) = 21.71$, p < 0.001
1voice2into	-0.93	-1.38 – -0.49	<0.001	
2voices1into	-0.15	-0.59 – 0.29	0.500	
2voices2into	-0.69	-1.13 – -0.25	0.002	
Group * Phonetic variability				$\chi^2(6) = 20.75$, p = .002
Percussion * 1voice2into	-0.33	-0.77 – 0.11	0.139	
Wind * 1voice2into	0.19	-0.22 – 0.60	0.360	
Percussion * 2voices1into	-0.20	-0.64 – 0.24	0.374	
Wind * 2voices1into	-0.07	-0.46 – 0.33	0.749	
Percussion * 2voices2into	-0.81	-1.24 – -0.38	<0.001	
Wind * 2voices2into	-0.43	-0.83 – -0.03	0.037	

Table 1: Summary of a logistic mixed-effects regression model. The reference levels are: Non-Musician for 'group', 1voice1into for 'phonetic variability'.