Rhythmic Abilities Correlate with L2 Prosody Imitation Abilities in Typologically Different Languages

Nia Cason, Muriel Marmursztejn, Mariapaola d'Imperio, Daniele Schön

HAL Id: hal-02349384
https://hal.science/hal-02349384

Original Article

Rhythmic Abilities Correlate with L2 Prosody Imitation Abilities in Typologically Different Languages

Nia Cason
Aix Marseille Univ, INSERM, INS, Inst Neurosci Syst, Marseille, France

Muriel Marmursztejn
Aix-Marseille Univ, CNRS, LPL, Laboratoire Parole et Langage, Aix-en-Provence, France

Mariapaola D'Imperio
Aix-Marseille Univ, CNRS, LPL, Laboratoire Parole et Langage, Aix-en-Provence, France

Daniele Schön [ID]
Aix Marseille Univ, INSERM, INS, Inst Neurosci Syst, Marseille, France

Abstract
While many studies have demonstrated the relationship between musical rhythm and speech prosody, this has been rarely addressed in the context of second language (L2) acquisition. Here, we investigated whether musical rhythmic skills and the production of L2 speech prosody are predictive of one another. We tested both musical and linguistic rhythmic competences of 23 native French speakers of L2 English. Participants completed perception and production music and language tests. In the prosody production test, sentences containing trisyllabic words with either a prominence on the first or on the second syllable were heard and had to be reproduced. Participants were less accurate in reproducing penultimate accent placement. Moreover, the accuracy in reproducing phonologically disfavored stress patterns was best predicted by rhythm production abilities. Our results show, for the first time, that better reproduction of musical rhythmic sequences is predictive of a more successful realization of unfamiliar L2 prosody, specifically in terms of stress-accent placement.

Keywords
Music, speech, foreign language, temporal patterns, stress patterns

Corresponding author:
Daniele Schön, Institut de Neurosciences des Systèmes, Faculté de Médecine, 27 bd Jean Moulin, Marseille, 13005, France.
Email: daniele.schon@univ-amu.fr

Introduction

Rhythm in both music and language organizes events in time. Rhythm perception has been thought to be an innate predisposition of human cognition that may be similarly processed across domains (e.g., Besson & Schön, 2001; Patel, 2014; Patel & Morgan, 2016). This is true for both "beat" (the perception of an underlying regular pulse) and "meter" (the perception of alternating weak and strong elements; London, 2004). Notably, the patterning of weak and strong elements is a fundamental characteristic of both speech and music, which could explain why rhythm in music has been compared to stress in language (Patel, 2006; Schön & Tillmann, 2015; Magne, Jordan, & Gordon, 2016). However, the links between rhythmic skills in music and speech have rarely been studied in the context of second language (L2) learning. The present study investigated the link between musical rhythmic competences and realization of L2 speech prosody, with the idea that better musical rhythmic skills would predict a more accurate realization of L2 prosody, specifically in terms of correct stress-accent placement (for the notion of stress-accent, see Beckman, 1986).

The perception of meter is crucial for language acquisition. During native language (L1) acquisition in infancy, metrical structures in speech serve as important cues in speech segmentation (Cutler & Butterfield, 1992; Jusczyk, Houston, & Newsome, 1999; Kuhl, 2004; Morgan & Saffran, 1995), which has been found to be facilitated by musical skills in childhood (François, Chobert, Besson, & Schön, 2013). One idea is that rhythmic entrainment (the coupling between the temporal structure of a stimulus and brain oscillatory activity) enhances the precision of auditory analyses and, as a consequence, enhances speech segmentation abilities (Hornickel & Kraus, 2013; Moreno et al., 2009). For instance, musical rhythmic abilities have been found to correlate with phonological abilities in preschoolers (Moritz, Yampolsky, Papadelis, Thomson, & Wolf, 2013), with reading ability in school-age children (Strait, Hornickel, & Kraus, 2011), and seem to be particularly relevant in decoding words requiring the use of linguistic stress (David, Wade-Woolley, Kirby, & Smithrim, 2007). Furthermore, metrical characteristics of L1 also affect rhythm perception (Iversen, Patel, & Oghushi, 2008; Kusumoto & Moreton, 1997) and may influence the rhythmic structure of instrumental music (Patel & Daniele, 2003), which suggests there to be a bidirectional influence of music and language temporal structures.

When turning to L2 learning, there is also evidence of a link between musical skills and L2 perception abilities (for reviews, see Chobert & Besson, 2013; Zeromskaite, 2014; Dittinger et al., 2016), including the detection of prosodic anomalies (Milovanov & Tervaniemi, 2011), prosodic pitch manipulations (Marquès, Moreno, Castro, & Besson, 2007), pitch contour (Zhao & Kuhl, 2015), speech imitation (Christiner & Reiterer, 2013), and phonological perception and production (Slevc & Myiake, 2006). These findings indicate that there may be shared resources between music and L2 acquisition. However, musical aptitude has largely been determined by musical melodic/pitch skills rather than rhythmic skills, which may be distinct (Phillips-Silver et al., 2011). Some evidence suggests that rhythmic skills may actually be a better predictor of L2 acquisition than pitch/melodic skills (Kempe, Bublitz, & Brooks, 2015). Indeed, recent work from Boll-Avetisyan, Bhatara, and Höhle (2017) found that musical rhythm aptitude, but not melody aptitude, was associated with rhythmic speech perception in native listeners.

Although links between L2 production abilities and musical rhythmic skills have rarely been demonstrated (Christiner & Reiterer, 2013; Christiner, Rüdegger, & Reiterer, 2018), rhythm and intonation have long been regarded as one of the greatest obstacles in L2 acquisition (Calbris & Montredon, 1975). This may be because L2 acquisition is strongly affected by the metrical characteristics of L1. Metrical "rules" that exist for L1 are used in the interpretation and segmentation of L2, even when these same rules do not apply (Cutler, Mehler, Norris, & Segui, 1986, 1992; Otake,

Hatano, Cutler, & Mehler, 1993). The language-specific use of metrical cues in French, for example, has been used to explain difficulties in encoding foreign stress contrasts that exist in Spanish (Dupoux, Pallier, Sebastian, & Mehler, 1997; Dupoux, Sebastian, Navarete, & Peperkamp, 2008), English (Kolinsky, Cuvelier, Goetry, Peretz, & Morais, 2009), and German (Schmidt-Kassow, Rothermich, Schwartze, & Kotz, 2011), which are attenuated by musical experience (Boll-Avetisyan, Bhatara, & Höhle, 2016).

In the present study, we investigated native French speakers with knowledge of English as an L2. The fundamental metrical differences between French and English can pose a challenge for native French speakers learning English. Indeed, French speakers find stress and pitch accent placement in English to be one of the most challenging points (see Capliez, 2011; Frost, 2008). Although French is considered to be a language in which lexical stress cues are absent (Dupoux et al., 1997, 2008; Cooper, Cutler, & Wales, 2002), it is characterized by the presence of stress at a phrasal level (the Accentual Phrase, or AP; see Jun & Fougeron, 2000; Post, 2000). In fact, while stress is a property of the word in English, it is a phrasal property in French. Namely, stress is located at an AP-final position in French (on the last full syllable of the AP), although an AP-initial pitch prominence is allowed in the case of an initial rise (Delattre, 1938; Hirst & Di Cristo, 1998; Fonagy, 1980; Vaissière, 1974). French speakers may therefore have a less-rich representation of speech meter compared to speakers of languages in which stress position is more variable, such as English (Di Cristo, 2003; Cutler & Carter, 1987; Tortel, 2009). Findings that native French speakers find it difficult to detect accentual contrasts that are not found in French support this idea (e.g., Dupoux et al., 2008; Kijak, 2009; Michelas, Frauenfelder, Schön, & Dufour, 2016; Bhatara, Boll-Avetisyan, Agus, Höhle, & Nazzi, 2016; Boll-Avetisyan et al., 2016).

To summarize, the link between rhythm in speech and in music is now quite well-established, but has yet to be investigated in the context of L2 prosody acquisition, especially in terms of speech production and imitation. Recent imitation studies have investigated production capabilities and individual differences in reproducing fine phonetic detail in prosody, such as tonal alignment (D'Imperio, Cavone, & Petrone, 2014) or syllable length (Cavone & D'Imperio, 2016). We thus investigated the relationship between rhythmic competences in music perception and production in relation to the ability for native French speakers to perceive and produce stress-accent placement in English words, specifically in word-initial and penultimate position. We hypothesized that musical rhythmic skills would be positively correlated with accuracy in the production of stress accent placement, and that this would be particularly evident for the penultimate stress, which is absent in the French system and is thus more difficult to realize. Our findings offer further support for a common processing of rhythm in language and music, and might also have practical implications for L2 learning techniques.

2　Materials And Methods

2.1 Participants

Twenty-three native French speakers studying at the science campus of Aix-Marseille University (16 female, age range 19–24 years, mean age 21 years) participated in the study. Participants had started learning English between the ages of 5 and 12 years (mean = 9.9 years; *SD* = 1.79) and had between 7 and 15 years of training in English (mean = 9.9 years; *SD* = 1.78). In other words, the level of English proficiency is that which French-speaking adults typically have gained from general school education. Twelve participants had received some musical training, which had started between the ages of 4 and 17 years (mean = 9.7 years, *SD* = 4.6; mean years of practice = 6.2, *SD* = 4.5). The remaining 11 participants had no musical practice (outside of the very basic school

curriculum). All participants gave their written formal consent and received a bonus of one point on their English grade for the semester for participating in this study. Participants completed both musical and linguistic tests, and the order of these tests was counterbalanced across participants.

2.2 Musical Tasks

Music perception. A subset of the Musical Ear Test (Wallentin, Nielsen, Friis-Olivarius, Vuust, & Vuust, 2010) was used to test melodic and rhythmic perception abilities. More precisely, we used the first 26 trials of each task instead of the entire 52 of the original Musical Ear Test. This was done so that all tasks could be completed in a single session. Because the Cronbach's α-consistency was very high for both melodic and rhythmic tests (0.95, reported in Wallentin et al., 2010), reducing the number of items should not make a major difference. For each trial, participants were asked to decide whether two short musical phrases (melodic or rhythmic) were the same or different. When different, two musical phrases differed from one another by a single element (one note for the melodic phrases, one rhythmic change for the rhythmic phrases).

Music production. *Rhythmic production*. Nine rhythmic sequences with a synthesized woodblock sound were played through loudspeakers. Each sequence consisted of three to seven sounds for a total of 46 sounds/taps (see Figure A1). Participants were orally asked to reproduce each rhythm once by tapping a wooden stick on a box right after listening to the sequence. No metronome cue was provided. A microphone was positioned next to the box and the participants' performance was recorded. A change-in-slope detection algorithm was used to extract the precise tapping times. We did not want to penalize globally slower or faster reproductions, so rhythmic production accuracy was measured by calculating the ratio of the productions of the 46 taps with those of the actual stimuli. For this, all intervals between two successive taps (t_n–t_{n1}) were computed, and each interval was divided by the interval that preceded it (t_{n1}–t_{n2}). The ratios for the participants' productions were then compared with those for the model stimuli (m), and the mean of the absolute value of the difference for all of the ratios obtained for all nine stimuli (N) was computed for each participant using the following formula:

$$\sum \mathrm{abs}\left[\left(t_n - t_{n1}\right)/\left(t_{n1} - t_{n2}\right) - \left(m_n - m_{n1}\right)/\left(m_{n1} - m_{n2}\right)\right]/N$$

This metric purposefully ignores the extent to which the tempo of the subject's drumming was similar to that of the stimulus. There was no reason to think that absolute tempo would be linked to stress placement, which is a relative phenomenon.

Melodic production. Participants were presented with a subset of stimuli from Lévêque and Schön (2013) through headphones, which comprised 10 stimuli of five isochronous notes each (see Figure A2). Participants were asked to reproduce each stimulus sequence once, immediately after hearing it. The melodic structure of the stimuli was tonal, which corresponds to what may be found in a popular song repertoire. The stimuli were sung by female or male voices and were presented accordingly to the participants' gender. The pitch range of the 50 pitches to be sung was adapted to adult tessitura and stimuli were presented one octave lower for male participants. PRAAT (Boersma & Weenink, 2009) was used to extract the median fundamental frequency of each note of the five-note vocal productions of the participants (discarding the transition glides). Then, the same procedure described for rhythmic production was used to compute the ratios of pitch frequency between successive intervals and to compare it to the model.

2.3 Linguistic Tasks

General English level assessment. An online test for the different levels of the Common European Framework of Reference for Languages was used (http://www.cambridgeenglish.org/test-your-english/adult-learners/). This consists of 25 multiple-choice questions presented in a written format that test vocabulary, grammar, and idiomatic expression knowledge to assess L2 English proficiency. Every correct item was given one point (score range 0–25).

Imitation of lexical stress in L2. This task was developed to assess the ability to correctly implement lexical stress in a foreign language. For this, we assessed the ability to implement a nuclear pitch accent associated with the second (penultimate) syllable of a trisyllabic word, which is a location that never receives primary stress in the participants' L1 (French). Participants heard sentences recorded by an American English speaker and were asked to immediately reproduce them. A script of each sentence was made available to reduce memory load. The 40 test sentences were organized in a simple Subject-Verb-Object mode and ended with an adverbial (one or two words, see Appendix B). The sentences contained either a high-frequency trisyllabic target word with a stress on the first syllable (e.g., *They met the MInister yesterday*) or on the penultimate syllable (e.g., *They got a new aPARtment last month*), which is a primary stress location that is rare in French (Hayes, 1995). The target word, selected from the CELEX database (Burnage et al., 1990), was the object noun in all Subject-Verb-Object sentences and thus bore the nuclear stress of the sentence, given that they were all-focus utterances. Frequency of lexical similarity (i.e., cognate words that are similar in both languages, such as "Minister/Ministre") was controlled across the two stress conditions (stress on syllable 1, stress on syllable 2); most target words were French–English cognates (32/40). Ten sentence fillers that did not contain any trisyllabic nouns were added. The sentences were presented in a pseudo-random order. Two native English speakers listened to the participants' recordings and used a binary rating system (1 correct; 0 incorrect) to evaluate whether the participants had succeeded or not in placing the nuclear accent on the correct syllable. The procedure was blind with respect to the results in the other tasks. When the two judges did not agree on a rating, the trial received a rating of 0.5. The two judges agreed on 82% of the words, with a κ-reliability score of 0.64. To further test the reliability of the judges, we added a third judge and obtained a Cohen κ-reliability score of 0.74, which is considered as a good inter-rater agreement (Fleiss, 1981).

The decision to use judges to decide whether accent production was correct or not was dictated by the absence of solidly established values that allow us to determine whether a syllable is stressed or not. Lexical stress is defined in terms of relative perceived prominence of a syllable in a multisyllabic word (Ladd, 1996/2008). Perceptually, stressed syllables are expected to be comparatively longer in duration, higher in pitch (when carrying a high or rising pitch accent), and perhaps louder than unstressed syllables (Lehiste & Lass, 1976). However, there are several obstacles to objective and automatic stress identification; acoustic correlates of stress and accent are highly language-specific (for a review, see Ortega-Llebaria & Prieto, 2011, for instance). Furthermore, stress production in L2 is subject to a variable amount of interference between L1 and L2 prosody. Hence, it is common practice to perform an auditory transcription of stress (see Michelas & D'Imperio, 2012, for French, and Cho & McQueen, 2005, for Dutch), given that the perception of prominence is categorical and susceptible to linguistic–perceptual judgment.

3 Results

First, we explored the results for each task and, when appropriate, compared performance against chance or across conditions or tasks. A descriptive correlation analysis was then performed to see whether the different tasks were related to each other. Because several variables were found to be
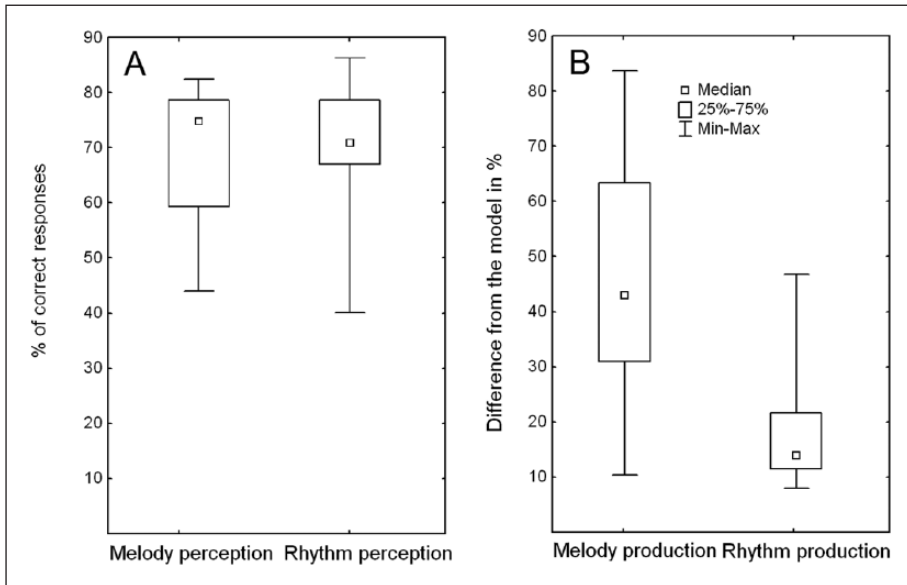
Figure 1. A: Boxplots of the % of correct responses for the music perception tasks.
B: Boxplots of the % difference from the stimuli for the music production tasks.
Min = minimum; max = maximum.

correlated, a factor analysis was conducted; this is a more advanced multivariate technique that allows the identification of variables that go well together and can be "summarized" in a small number of factors. Finally, a regression analysis was performed, which allowed us to test the hypothesis that performance of stress placement in ultimate and penultimate positions is best predicted by language proficiency or musical (melodic or rhythmic) skills.

3.1 Music and Language Tasks

There were no significant differences between the % of correct responses in the melody and rhythm perception tasks (Figure 1A; Wilcoxon, $z = 0.5$, $p = 0.6$). However, the population dispersion was smaller in the rhythm perception task, which indicates a more homogenous level of performance (although there was one participant with a very poor rhythmic performance). Participants were able to perform the melody and rhythm perception tasks well above chance level (one sample t-test, always $p < 0.001$). For music production tasks (Figure 1B), a more homogenous level of performance was also seen for rhythm compared to melody. Note that, although data are represented in terms of % distance from the model, it is impossible to directly compare melody and rhythm results due to the difference in units (Hertz vs. milliseconds, respectively). We also computed the α-Cronbach measure of reliability (Cronbach, 1951) to assess whether our measures were robust. Both melody and rhythm production showed a very high internal consistency (α = 0.93 and 0.86, respectively).

Scores for the English proficiency test (Figure 2A) showed a rather widespread distribution with a median of 13/25 (range 7–20). Thus, the population we tested did not include participants with beginner or expert English proficiency levels. For the speech imitation task (Figure 2B), participants were better able to correctly reproduce stress when it occurred on the first compared to the second syllable (Wilcoxon, $z = 3.2$, $p = 0.001$).
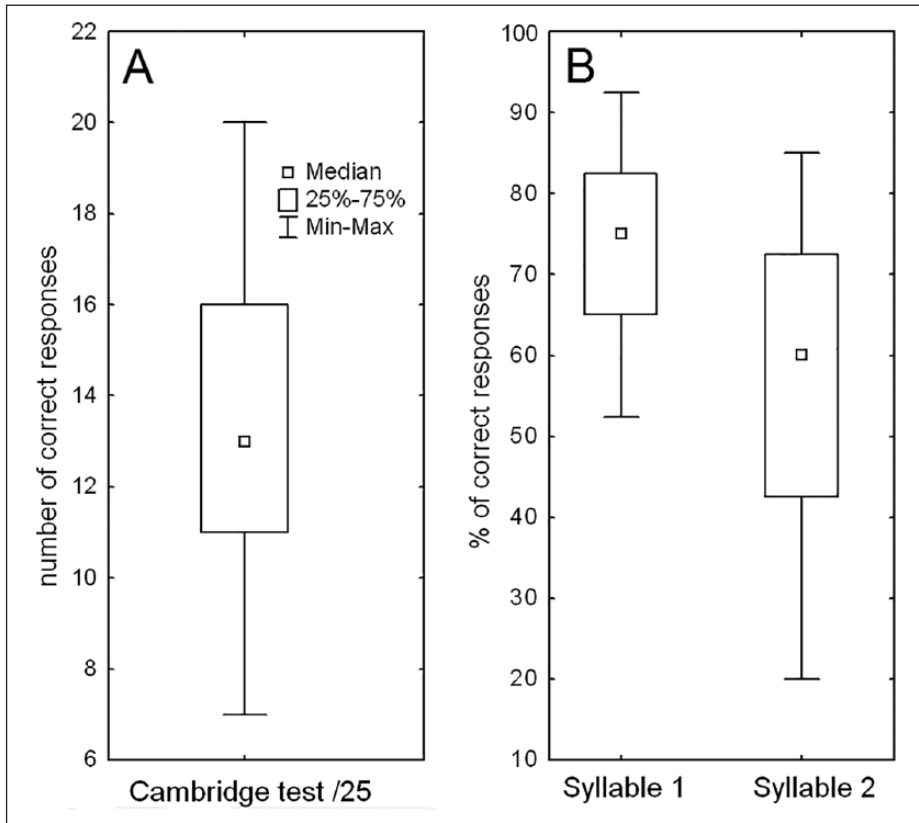
Figure 2. A: Boxplot of the number of correct responses for the Cambridge test.
B: Boxplots of the % of correct responses for the imitation task when the prominence was on the first or in the second syllable of the trisyllabic words.
Min = minimum; max = maximum.

3.2 Correlation Analysis

A Spearman's correlation analysis (based on ranks) was performed to test the strength of relationships between variables. Correlations between all the musical and language tasks are provided in Table 1. An overview of significant values (in bold, $p < 0.05$, one-tailed, uncorrected) shows that accuracy in the imitation task was correlated with English proficiency level. Significant correlations were also found between the musical perception tasks (melody and rhythm) and between the musical production tasks (melody and rhythm). Rhythm production was the only musical dimension that was correlated with imitation accuracy, when the stress had to be placed on the second syllable. A post-hoc power analysis was computed using the software G Power 3.1.9.2 given α (0.05), sample size and effect size (r) and showed that significant correlations always achieved a power greater than 0.7 (range 0.7–0.9). The interdependence between these variables was further tested using a factor analysis.

3.3 Factor Analysis

The factor analysis included results from the language tasks (English proficiency level and imitation) and the music perception and production tasks (rhythm and melody). Preliminary testing

Table 1. Spearman's correlation coefficients between the dependent variables.

	L2 proficiency level	Melody perception	Rhythm perception	Melody production	Rhythm production	Syllable 1 Imitation
Melody perception	−0.03					
Rhythm perception	0.23	**0.62** **(0.35, 1)**				
Melody production	−0.23	**−0.45** **(−0.8, −0.1)**	−0.35			
Rhythm production	−0.35	−0.36	**−0.5** **(−0.9, −0.15)**	0.14		
Syllable 1 imitation	**0.59** **(0.34, 1)**	0.06	0.21	0.04	−0.1	
Syllable 2 imitation	**0.47** **(0.1, 0.8)**	0.33	0.16	−0.12	**−0.53** **(−0.9, −0.15)**	0.39

Bold represents significant correlations. Values in parentheses indicate the 95% confidence interval.

Table 2. Varimax rotated factor loadings for the language and music variables, using the option "Blank" ($<$|0.40|).

	Factor 1	Factor 2
	(−3.16)	(−1.71)
L2 proficiency level	−0.54	**0.70**
Melody perception	**−0.76**	−0.47
Rhythm perception	**−0.83**	
Melody production	−0.49	
Rhythm production	**−0.85**	
Syllable 1 imitation		**0.79**
Syllable 2 imitation	**−0.71**	0.40
Explained variance	45%	24%

Bold represents significant correlations.
L2 = second language.

showed that our model was satisfactorily adequate; The Kaiser-Meyer-Olkin index measuring the sampling adequacy gave a value greater than 0.6. Finally, following two different methods to estimate the number of factors (using the software package FACTOR, Unrestricted Factor Analysis 9.2 by Urbano Lorenzo-Seva and Pere J. Ferrando) and an eigenvalue criterion $\geqslant 1$, two factors were extracted that explained 69% of the variance (Table 2).

The first factor showed high factor loadings (i.e., correlation coefficients between variables and factors) for melody and rhythm perception, rhythm production, and imitation accuracy of the syllable with disfavored lexical stress. Thus, this first factor can be interpreted as describing musical abilities. The second factor showed high factor loadings for English proficiency scores and imitation accuracy of words with initial-syllable stress. It can thus be interpreted as a factor that describes more specific linguistic abilities.

3.4 Regression Analyses

In the regression analyses, the two outcomes of the imitation task (imitation of syllable 1 and of syllable 2) were considered as dependent variables in separate analyses. The outcome of the factor

analysis was used to select the most appropriate predictors, as follows: L2 proficiency and melody perception for syllable 1 and rhythm production and rhythm and melody perception for syllable 2. When considering imitation accuracy of the first syllable, L2 proficiency significantly predicted the outcome ($\beta = 0.61$, $F = 12.8$, $p = 0.002$). When considering imitation accuracy of the penultimate syllable, only rhythm production significantly predicted the outcome ($\beta = 0.81$, $F = 7.8$, $p = 0.01$), whereas rhythm and melody perception did not reach significance (rhythm: $\beta = 0.23$, $F < 1$; melody: $\beta = 0.08$, $F < 1$). We computed f^2 Cohen values from the R^2 values of the regression and used these as effect sizes, together with α (0.05) and sample size (23) to compute post-hoc achieved power using G*Power 3.1.9.2. Although the post-hoc power analysis showed that significant regressors achieved a power > 0.8, results of the regression analysis should nonetheless be treated with caution due to the rather limited sample size.

4   Discussion

One of the greatest challenges in L2 acquisition is being able to reproduce prosodic structure, particularly rhythm and intonation (Calbris & Montredon, 1975). We used a novel task to investigate the links between musical skills and the ability to imitate speech prosody in L2, which is an element of speech that is crucial for decoding the speech signal. French participants were more able to imitate accent placement for word-initial than for penultimate stress location in English L2. This was expected given the constraint against moving the primary stress location away from the edges of the AP in French. Most importantly, correct placement of the nuclear accent was well predicted by rhythm production scores. More precisely, participants with better rhythm reproduction ability were more able to reproduce penultimate stress by placement of a correct nuclear accent. Whereas previous studies have investigated the possible links between rhythm perception and language abilities in the context of L1 perception (Hausen et al., 2013), ours is the first study to show a link between musical rhythm production and the ability to correctly imitate stress-accent placement in an L2 that is characterized by a typologically different prosodic system. Our results suggest that rhythmic training may possibly have a facilitatory effect in learning appropriate L2 prosody (and vice versa; see Bhatara, Yeung, & Nazzi, 2015, for a similar conclusion on the link between foreign language experience L2 and rhythm perception).

4.1 L2 Acquisition and Musical Rhythmic Skills

Several studies have found a relationship between musicianship and L2 skills (e.g., Lee & Hung, 2008). However, studies investigating the relationship between specific musical skills and L2 have primarily focused on perception and general L2 proficiency (Bhatara et al., 2015; Boll-Avetisyan et al., 2016), or have looked at links between general musical experience and L2 rhythm (Bhatara et al., 2016). Other studies have focused on the musical dimension of pitch processing and have not systematically attempted to tease-apart pitch and rhythmic abilities (Posedel, Emery, Souza, & Fountain, 2012; Milovanov, Huotilainen, Välimäki, Esquef, & Tervaniemi, 2008; Slevc & Miyake, 2006). Not only does this mean that there are relatively few studies concerned with temporal processing, which is crucial for language and speech comprehension (Goswami, 2011), but it is also questionable how pitch processing in music can be compared to pitch processing in speech (such as in tonal and intonational languages). For instance, a small inaccuracy in pitch-level production is extremely salient in music (e.g., one-third of a tone), but goes unnoticed in speech. Saying this, musical relative pitch skills may be relevant when studying factors that may impact the acquisition of tonal languages; notably, non-native musicians more accurately perceive Mandarin tones compared to nonmusicians (Lee & Hung, 2008; Gottfried & Riester, 2000).

Regarding the temporal aspects of speech, musical expertise has been found to have an effect on the discrimination of consonant and vowel duration in both L1 and L2 (Sadakata & Sekiyama, 2011) and the identification of segmental and tonal contrasts in L2 (Marie et al., 2011). In a longitudinal study, François et al. (2013) found that music training predicts more efficient speech segmentation skills in an artificial language in children. Overall, well-developed rhythmic abilities could translate into a better detection of speech regularities (Kraus, Strait, & Parbery-Clark, 2012). Additionally, music-L2 links are most probably bidirectional (Bidelman, Hutka, & Moreno, 2013). For instance, an enhanced music rhythm perception has been found for those who have mastered an L2 (Roncaglia-Denissen, Roor, Chen, & Sadakata, 2016); this is particularly true when the L2 is rhythmically different relatively to the L1, an effect that cannot be simply explained by exposure to more complex musical rhythms (as in Turkish L1 speakers; Roncaglia-Denissen et al., 2016). This relationship between musical rhythm and speech perception could rely on shared cognitive functions (such as working memory), as well as on the ability to entrain to multiplexed temporal scales, from the millisecond to second levels (Tierney & Kraus, 2015; Doelling & Poeppel, 2015; Schön & Tillmann, 2015). Future studies could implement a longitudinal design to demonstrate a causal effect of specific features of musical training (here, rhythm) on other features of L2 acquisition (here, accent placement).

As noted above, previous findings have been mainly concerned with L2 perception, and not L2 production. One of the only studies to have investigated the effect of music skills on L2 production showed that musical competence has an impact on phonological perception and production (Slevc & Miyake, 2006). However, this study only tested melodic competence, and in terms of L2 production it only tested segmental phonology. The present study is the first to investigate the effects of musical expertise on lexical accent placement in L2 production, and to investigate its relationship with both melodic and rhythmic perception and production skills. Similarly to the present study, participants in Slevc and Miyake's (2006) study (Japanese adults learning English) had an existing L2 knowledge. This means that the L2 was, nonetheless, familiar to some degree, and correct stress placement in the current study was thus also due to linguistic training. A distinction can therefore be made with studies that have looked at the effect of musical expertise on the production of a completely unknown language (Christiner & Reiterer, 2013, 2015; Christiner et al., 2018). Although "sense of rhythm" was considered to play a role in the ability to imitate an unknown language (Hindi), the authors of the 2013 study found that 66% of the speech imitation variance could be explained by working memory capacity, possibly due to intrinsic high memory load when imitating a completely unknown language. In this respect, an advantage of working with a known L2, as in our case, is that it is easy to provide participants with a printed version of the heard sentences to be reproduced, thus preventing a memory load bias.

Interestingly, we did not find any relationship between melodic performance and L2 accent placement imitation success. This finding fits with the growing idea that rhythm plays a key role in speech perception and production, and supports findings that rhythm—and not melody—is most effective in language/speech therapy (Overy, Nicolson, Fawcett, & Clarke, 2003; Goswami, 2011; Stahl, Kotz, Henseler, Turner, & Geyer, 2011; Flaugnacco et al., 2014). However, an analysis of more global pronunciation skills (as in Christiner & Reiterer, 2015) might have revealed a relationship with melodic aptitude (e.g., see Milovanov et al., 2008). Future studies could more closely examine the relative contributions of different musical skills to different L2 skills, and how these contributions differ according to the L2 being learned; and also in relation to the properties of one's L1 (see Roncaglia-Denissen et al., 2016).

The fact that L2 English proficiency (measured using the standard Cambridge test) was also predictive of a correct prosodic production may simply reflect the fact that a more thorough knowledge of the L2 implies heightened exposure, hence familiarity, with its prosody (see also

D'Imperio & German, 2015). Even more interesting is that, while we would expect English proficiency to actually be the *best* predictor of L2 prosody production, the best predictor of the performance on the most difficult items (i.e., penultimate stress) was rhythm production score. Future studies should investigate the relationship between musical rhythmic skills and prosody production using a known and an unknown L2 language so as to evaluate the effect of negative transfer. Indeed, previous knowledge of L1 might interfere with new learning (L2), which may result in establishing bad pronunciation habits that may not be present when testing another, unknown L2 (Odlin, 1989). It would also be interesting to compare imitation abilities with L2 that vary in their linguistic distance to L1 (such as in D'Imperio & German, 2015, in which two typologically different prosodic systems were compared). For instance, in our study, some items were lexically similar in L1 and L2, but had a different prosody (e.g., **MI**nister/min**IS**tre). Musical rhythmic skills might therefore play a different role when there must be a process of reconfiguration of the stress position with respect to L1 compared to when there is no lexical competition and no need for stress pattern reconfiguration.

4.2 Theoretical and Practical Implications: A Common Processing of Meter

The implications of these findings are theoretically and practically relevant. From a theoretical standpoint, our results show that metrical rhythmic patterns may be processed similarly for language and music. Namely, participants who were more accurately able to reproduce musical rhythms also showed an advantage in the realization of syllables with a disfavored stress location. This finding extends and refines previous studies showing that musical expertise is correlated with L2 skills (see Chobert & Besson, 2013, and Zeromskaite, 2014, for reviews), and suggests that musical temporal aptitude may be linked to specific aspects of foreign language acquisition.

From a practical standpoint, these findings are interesting when considering the value of music-based teaching techniques for foreign language teachers. Besides a general motivational benefit, different features of music could be used to reinforce prosodic representations of the foreign language being taught. Considering previous findings in L1 acquisition (Cason, Hidalgo, & Schön, 2015) and in children with reading impairments (Bhide et al., 2013; Flaugnacco et al., 2015), rhythmic musical training can also be predicted to impact positively on L2 phonological perception and production, too.

It has been shown that musical competence is not uniquely the result of music training, but of many other cognitive and social factors (Swaminathan & Schellenberg, 2018). However, musicians outperform nonmusicians in several musical skills, including temporal rhythmic skills (Rammsayer & Altenmüller, 2006; Rammsayer, Buttkus, & Altenmüller, 2012; Schaal, Banissy, & Lange, 2015; Wollman & Morillon, 2018). Thus, music training or music-based L2 learning, by improving temporal processing at different levels, may in turn facilitate the learning of segmental and suprasegmental features of an L2.

Importantly, different aspects of musical training may make different contributions to L2 acquisition, depending on the language. For instance, we might expect to find similar results for Turkish L1 speakers learning English L2, because Turkish has a word-final stress (Kabak & Vogel, 2001). Similarly, several other L2 skills may instead benefit from pitch training, particularly in languages in which pitch is more important (e.g., Mandarin). Finally, due to the metrical nature of French prosody, French speakers may have less-rich representations of speech meter than speakers of more prosodically variable languages, and may thus benefit more from a rhythmic music-based teaching approach than might speakers of other languages. Future studies could investigate the relationship of musical rhythmic skills and L2 prosodic skills in speakers whose L1 is more prosodically variable. To conclude, while further research is required to specify the L2-specific links

with musical abilities, this study demonstrates a strong relationship between the perception and production ability of L2 meter and musical rhythmic production skills.

ORCID iD

Daniele Schön [iD] https://orcid.org/0000-0003-4472-4150

References

Beckman, M. E. (1986). *Stress and non-stress accent*. Dordrecht, Netherlands: Foris Publications.

Besson, M., & Schön, D. (2001). Comparison between language and music. *Annals of the New York Academy of Sciences*, *930*(1), 232–258.

Bhatara, A., Boll-Avetisyan, N., Agus, T., Höhle, B., & Nazzi, T. (2016). Language experience affects grouping of musical instrument sounds. *Cognitive Science*, *40*(7), 1816–1830.

Bhatara, A., Yeung, H. H., & Nazzi, T. (2015). Foreign language learning in French speakers is associated with rhythm perception, but not with melody perception. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(2), 277–282.

Bhide, A., Power, A., & Goswami, U. (2013). A rhythmic musical intervention for poor readers: A comparison of efficacy with a letter-based intervention. *Mind, Brain, and Education*, *7*(2), 113–123.

Bidelman, G. M., Hutka, S., & Moreno, S. (2013). Tone language speakers and musicians share enhanced perceptual and cognitive abilities for musical pitch: Evidence for bidirectionality between the domains of language and music. *PLoS ONE*, *8*(4). Retrieved from https://doi.org/10.1371/journal.pone.0060676

Boersma, P., & Weenink, D. (2009). PRAAT: Doing phonetics by computer (version 5.1.34) [Computer software]. Retrieved from http://www.praat.org

Boll-Avetisyan, N., Bhatara, A., & Höhle, B. (2017). Effects of musicality on the perception of rhythmic structure in speech. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *8*(1), 9.

Boll-Avetisyan, N., Bhatara, A., Unger, A., Nazzi, T., & Höhle, B. (2016). Effects of experience with L2 and music on rhythmic grouping by French listeners. *Bilingualism: Language and Cognition*, *19*(5), 971–986.

Burnage, G. (1990). *CELEX: A guide for users*. Nijmegen, Netherlands: CELEX Centre for Lexical Information. Retrieved from: http://groups.linguistics.northwestern.edu/speech_comm_group/documents/CELEX/Full%20Celex%20User%20Guide.pdf

Calbris, G., & Montredon, J. (1975). *Approche rythmique intonative du français langue étrangère* (Vol. 1). Paris, France: Clé International.

Capliez, M. (2011). Typologie des erreurs de production d'anglais des francophones: Segments vs. suprasegments. Recherche et pratiques pédagogiques en langues de spécialité. *Cahiers de l'Apliut*, *30*(3), 44–60.

Cason, N., Hidalgo, C., & Schön, D. (2015). Rhythmic priming enhances speech production abilities: Evidence from prelingually deaf children. *Neuropsychology*, *29*(1), 102.

Cavone, R., & D'Imperio, M. (2016). Poster presented at LabPhon 15: *L1 use predicts imitation of metrical features in a typologically different L2*. Cornell University, Ithaca, NY. Retrieved from http://labphon15.labphon.org/long_abstracts/LabPhon15_Revised_abstract_111.pdf

Cho, T., & McQueen, J. M. (2005). Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress, *Journal of Phonetics*, *33*, 121–157.

Chobert, J., & Besson, M. (2013). Musical expertise and second language learning. *Brain Sciences*, *3*(2), 923–940.

Christiner, M., & Reiterer, S. M. (2013). Song and speech: Examining the link between singing talent and speech imitation ability. *Frontiers in Psychology*, *21*(4), 874.

Christiner, M., & Reiterer, S. M. (2015). A Mozart is not a Pavarotti: Singers outperform instrumentalists on foreign accent imitation. *Frontiers in Human Neuroscience*, *9*, 482.

Christiner, M., Rüdegger, S., & Reiterer, S. M. (2018). Sing Chinese and tap Tagalog? Predicting individual differences in musical and phonetic aptitude using language families differing by sound-typology. *International Journal of Multilingualism*, *15*(4), 455–471.

Cooper, N., Cutler, A., & Wales, R. (2002). Constraints of lexical stress on lexical access in English: Evidence from native and non-native listeners. *Language and Speech*, *45*(3), 207–228.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334.

Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language*, *31*, 218–236.

Cutler, A., & Carter, D. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language*, *2*, 133–142.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language*, *25*, 385–400.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1992). The monolingual nature of speech segmentation by bilinguals. *Cognitive Psychology*, *24*(3), 381–410.

David, D., Wade-Woolley, L., Kirby, J. R., & Smithrim, K. (2007). Rhythm and reading development in school-age children: A longitudinal study. *Journal of Research in Reading*, *30*(2), 169–183.

Delattre, P. (1938). L'accent final en français: Accent d'intensité, accent de hauteur, accent de durée. *French Review*, *12*, 141–145.

Di Cristo, A. (2003). De la métrique et du rythme de la parole ordinaire: L'exemple du français. *Semen: Revue de Sémio-Linguistique des Textes et Discours*, *16*. Retrieved from https://journals.openedition.org/semen/2944

D'Imperio, M., Cavone, R., & Petrone, C. (2014). Phonetic and phonological imitation of intonation in two varieties of Italian. *Frontiers in Psychology*. Retrieved from https://www.doi.org/10.3389/fpsyg.2014.01226

D'Imperio, M., & German, J. S. (2015). Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015): *Phonetic detail and the role of exposure in dialect imitation*. International Phonetic Association. Retrieved from https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS1009.pdf

Dittinger, E., Barbaroux, M., D'Imperio, M., Jäncke, L., Elmer, S., & Besson, M. (2016). Professional music training and novel word learning: From faster semantic encoding to longer-lasting word representations. *Journal of Cognitive Neuroscience*, *28*, 1584–1602.

Doelling, K. B., & Poeppel, D. (2015). Cortical entrainment to music and its modulation by expertise. *Proceedings of the National Academy of Sciences*, *112*(45), E6233–E6242.

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

Dupoux, E., Sebastian, N., Navarete, E., & Peperkamp, S. (2008). Persistent stress deafness: The case of French learners of Spanish. *Cognition*, *106*(2), 682–706.

Flaugnacco, E., Lopez, L., Terribili, C., Montico, M., Zoia, S., & Schön, D. (2015). Music training increases phonological awareness and reading skills in developmental dyslexia: A randomized control trial. *PLoS One*, *10*(9), e0138715.

Flaugnacco, E., Lopez, L., Terribili, C., Zoia, S., Buda, S., Tilli, S., … Schön, D. (2014). Rhythm perception and production predict reading abilities in developmental dyslexia. *Frontiers in Human Neuroscience*, *8*. Retrieved from https://www.doi.org/10.3389/fnhum.2014.00392

Fleiss, J. L. (1981). *Statistical methods for rates and proportions*, 2nd ed (pp. 38–46). New York, NY: John Wiley.

Fonagy, I. (1980). L'accent français, accent probabilitaire. *Studia Phonetica*, *15*, 123–133.

François, C., Chobert, J., Besson, M., & Schön, D. (2013). Music training for the development of speech segmentation. *Cerebral Cortex*, *23*(9), 2038–2043.

Frost, D. (2008). *The stress site: L'accent lexical, l'anglais de spécialité et l'oral, la conception d'un outil d'apprentissage médiatisé* (Doctoral dissertation). Bordeaux, France: Université de Bordeaux.

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Science*, *15*(1), 3–10.

Gottfried, T. L., & Riester, D. (2000). Relation of pitch glide perception and Mandarin tone identification. *The Journal of the Acoustical Society of America*, *108*(5), 2604.

Hausen, M., Torppa, R., Salmela, V., Vainio, M., & Sarkamo, T. (2013). Music and speech prosody: A common rhythm. *Frontiers in Psychology*. Retrieved from https://www.doi.org/10.3389/fpsyg.2013.00566

Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. Chicago, IL: University of Chicago Press.

Hirst, D., & Di Cristo, A. (1998). *Intonation systems: A survey of twenty languages*. New York, NY: Cambridge University Press.

Hornickel, J., & Kraus, N. (2013). Unstable representation of sound: A biological marker of dyslexia. *Journal of Neuroscience*, *33*(8), 3500–3504.

Iversen, J. R., Patel, A. D., & Ohgushi, K. (2008). Perception of rhythmic grouping depends on auditory experience. *The Journal of the Acoustical Society of America*, *124*, 2263–2271.

Jun, S. A., & Fougeron, C. (2000). A phonological model of French intonation. In A. Botinis (Ed.), *Intonation: Analysis, modeling and technology*. Dordrecht, Netherlands: Kluwer.

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, *39*, 159–207.

Kabak, B., & Vogel, I. (2001). The phonological word and stress assignment in Turkish. *Phonology*, *18*, 315–360.

Kempe, V., Bublitz, D., & Brooks, P. J. (2015). Musical ability and non-native speech-sound processing are linked through sensitivity to pitch and spectral information. *British Journal of Psychology*, *106*, 349–366.

Kijak, A. (2009). How stressful is L2 stress? A cross linguistic study of L2 perception and production of metrical systems (Doctoral dissertation). *Universiteit Utrecht*. Retrieved from https://www.lotpublications.nl/Documents/214_fulltext.pdf

Kolinsky, R., Cuvelier, H., Goetry, V., Peretz, I., & Morais, J. (2009). Music training facilitates lexical stress processing. *Music Perception*, *26*(3), 235–246.

Kraus, N., Strait, D. L., & Parbery-Clark, A. (2012). Cognitive factors shape brain networks for auditory skills: Spotlight on auditory working memory. *Annals of the New York Academy of Sciences*, *1252*, 100–107.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, *5*, 831–843.

Kusumoto, K., & Moreton, E. (1997). *Native language determines the parsing of nonlinguistic rhythmic stimuli* (Doctoral dissertation). Melville, NY: Acoustical Society of America.

Ladd, D. R. (1996/2008). *Intonational phonology* (2nd ed.). Cambridge, UK: Cambridge University Press.

Lee, C. Y., & Hung, T. H. (2008). Identification of Mandarin tones by English-speaking musicians and non-musicians. *The Journal of the Acoustical Society of America*, *124*, 3235–3248.

Lehiste, I., & Lass, N. J. (1976). Suprasegmental features of speech. In N. J. Lass (Ed.), *Contemporary issues in experimental phonetics* (pp. 225–239). New York, NY: Academic Press.

Lévêque, Y., & Schön, D. (2013). Listening to the human voice alters sensorimotor brain rhythms. *PLoS One*, *8*(11), e80659.

London, J. (2004). *Hearing in time: Psychological aspects of musical meter*. New York, NY: Oxford University Press.

Magne, C., Jordan, D. K., & Gordon, R. L. (2016). Speech rhythm sensitivity and musical aptitude: ERPs and individual differences. *Brain & Language*, *153*, 13–19.

Marie, C., Delogu, F., Lampis, G., Belardinelli, M. O., & Besson, M. (2011). Influence of musical expertise on segmental and tonal processing in Mandarin Chinese. *Journal of Cognitive Neuroscience*, *23*(10), 2701–2715.

Marques, C., Moreno, S., Castro, S., & Besson, M. (2007). Musicians detect pitch violation in a foreign language better than nonmusicians: Behavioral and electrophysiological evidence. *Journal of Cognitive Neuroscience*, *19*(9), 1453–1463.

Michelas, A., & D'Imperio, M. (2012). When syntax meets prosody: Tonal and duration variability in French accentual phrases. *Journal of Phonetics*, *40*(6), 816–829.

Michelas, A., Frauenfelder, U. H., Schön, D., & Dufour, S. (2016). How deaf are French speakers to stress? *The Journal of the Acoustical Society of America*, *139*(3): 1333–1342.

Milovanov, R., Huotilainen, M., Välimäki, V., Esquef, P. A., & Tervaniemi, M. (2008). Musical aptitude and second language pronunciation skills in school-aged children: Neural and behavioral evidence. *Brain Research*, *1194*, 81–89.

Milovanov, R., & Tervaniemi, M. (2011). The interplay between musical and linguistic aptitudes: A review. *Frontiers in Psychology*, *2*. Retrieved from https://www.doi.org/10.3389/fpsyg.2011.00321

Moreno, S., Marques, C., Santos, A., Santos, M. C., & Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex*, *19*(3), 712–723.

Morgan, J. L., & Saffran, J. R. (1995). Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. *Child Development*, *66*, 911–936.

Moritz, C., Yampolsky, S., Papadelis, G., Thomson, J., & Wolf, M. (2013). Links between early rhythm skills, musical training, and phonological awareness. *Reading and Writing*, *26*, 739–769.

Odlin, T. (1989). Language transfer: Cross-linguistic influence in language learning. New York, NY: Cambridge University Press.

Ortega-Llebaria, M., & Prieto, P. (2011). Acoustic correlates of stress in Central Catalan and Castilian Spanish. *Language and Speech*, *54*(1), 73–97.

Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language*, *32*, 258–278.

Overy, K., Nicolson, R., Fawcett, A., & Clarke, E. (2003). Dyslexia and music: Measuring musical timing skills. *Dyslexia*, *9*, 18–36.

Patel, A. D. (2006). Musical rhythm, linguistic rhythm, and human evolution. *Music Perception*, *24*, 99–104.

Patel, A. D. (2014). The evolutionary biology of musical rhythm: Was Darwin wrong? *PLoS Biology*, *12*(3): e1001821.

Patel, A. D., & Daniele, J. R. (2003). An empirical comparison of rhythm in language and music. *Cognition*, *87*(1): B35–B45.

Patel, A. D., & Morgan, E. (2016). Exploring cognitive relations between prediction in language and music. *Cognitive Science*. Retrieved from https://www.doi.org/10.1111/cogs.12411

Phillips-Silver, J., Toiviainen, P., Gosselin, N., Piché, O., Nozaradan, S., Palmer, C., & Peretz, I. (2011). Born to dance but beat deaf: A new form of congenital amusia. *Neuropsychologia*, *49*, 961–969.

Posedel, J., Emery, L., Souza, B., & Fountain, C. (2012). Pitch perception, working memory, and second-language phonological production. *Psychology of Music*, *40*, 508–517.

Post, B. (2000). *Tonal and phrasal structures in French intonation* (Doctoral dissertation). University of Nijmegen. The Hague, Netherlands: Holland Academic Graphics.

Rammsayer, T., & Altenmüller, E. (2006). Temporal information processing in musicians and nonmusicians. *Music Perception: An Interdisciplinary Journal*, *24*(1), 37–48.

Rammsayer, T. H., Buttkus, F., & Altenmüller, E. (2012). Musicians do better than nonmusicians in both auditory and visual timing tasks. *Music Perception: An Interdisciplinary Journal*, *30*(1), 85–96.

Roncaglia-Denissen, M. P., Roor, D. A., Chen, A., & Sadakata, M. (2016). The enhanced musical rhythmic perception in second language learners. *Frontiers in Human Neuroscience*, *10*, 288.

Sadakata, M., & Sekiyama, K. (2011). Enhanced perception of various linguistic features by musicians: A cross-linguistic study. *Acta Psychologica*, *138*, 1–10.

Schaal, N. K., Banissy, M. J., & Lange, K. (2015). The rhythm span task: Comparing memory capacity for musical rhythms in musicians and non-musicians. *Journal of New Music Research*, *44*(1), 3–10.

Schmidt-Kassow, M., Rothermich, K., Schwartze, M., & 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.

Schön, D., & Tillmann, B. (2015). Short-and long-term rhythmic interventions: Perspectives for language rehabilitation. *Annals of the New York Academy of Sciences*, *1337*, 32–39.

Slevc, R., & Miyake, A. (2006). Individual differences in second-language proficiency: Does musical ability matter? *Psychological Science*, *17*, 675–681.

Stahl, B., Kotz, S. A., Henseler, I., Turner, R., & Geyer, S. (2011). Rhythm in disguise: Why singing *may* not hold the key to recovery from aphasia. *Brain*, *134*, 3083–3093.

Strait, D. L., Hornickel, J., & Kraus, N. (2011). Subcortical processing of speech regularities underlies reading and music aptitude in children. *Behavioral and Brain Functions*, *7*(1), 44.

Swaminathan, S., & Schellenberg, E. G. (2018). Musical Competence is Predicted by Music Training, Cognitive Abilities, and Personality. *Scientific reports*, *8*(1), 9223.

Tierney, A., & Kraus, N. (2015). Neural entrainment to the rhythmic structure of music. *Journal of Cognitive Neuroscience*, *27*(2): 400–408.

Tortel, A. (2009). *Evaluation qualitative de la prosodie d'apprenants français: Apports de paramétrisations prosodiques* (Doctoral dissertation). Marseille, France: Université Aix-Marseille I.

Vaissière, J. (1974). On French prosody. *Quarterly Progress Report, Research Laboratory of Electronics, Massachusetts Institute of Technology*, *114*, 212–223.

Wallentin, M., Nielsen, A. H., Friis-Olivarius, M., Vuust, C., & Vuust, P. (2010). The musical ear test, a new reliable test for measuring musical competence. *Learning and Individual Differences*, *20*(3), 188–196.

Wollman, I., & Morillon, B. (2018). Organizational principles of multidimensional predictions in human auditory attention. *Scientific Reports*, *8*(1), 13466.

Zhao, T. C., & Kuhl, P. K. (2015). Effect of musical experience on learning lexical tone categories. *The Journal of the Acoustical Society of America*, *137*, 1452–1463.

Zeromskaite, I. (2014). The potential role of music in second language learning: A review article. *Journal of European Psychology Students*, *5*(3), 78–88.

Appendix A

Figure A1. Music notation of the stimuli used in the Rhythm Production test.

Figure A2. Music notation of the stimuli used in the Melody Production test.

Appendix B. Linguistic corpus for the imitation task.

Utterances containing a trisyllabic word with a stress on the first syllable	Utterances containing a trisyllabic word with a stress on the second syllable
They spotted the criminal yesterday.	I went to the casino yesterday.
Claire's going to the festival today.	They've felt some attraction lately.
Lisa will change the battery tomorrow.	Justin gave some perspective yesterday.
They look for happiness everyday.	They spoke to the detective yesterday.
Mary will open a gallery next year.	We will create a new republic soon.
James spoke to the citizen last night.	We have an appointment tomorrow.
They like the atmosphere very much.	Sara hired a new assistant last month.
Ken will open a factory next month.	We like potatoes a lot.
Peter will write an article tomorrow.	They lost their objective in the end.
They had an accident last year.	Jason's felt some improvement lately.
We've found a new element lately.	They got a new apartment last month.
They will take some holiday next month.	They chose the other collection yesterday.
They talked about politics yesterday.	They hired a new director last month.
We did the exercise last night.	They had a good impression after all.
Claire will go to Africa next year.	Laura found a solution last week.
Mike's changed his attitude lately.	They hired a new professor last year.
They went to the hospital yesterday.	They will buy a new computer next month.
They met the minister yesterday.	They started the production last year.
They changed their policy last year.	We will elect a new committee next month.
They met with the government last month.	They will call the inspector tomorrow.

"Fillers"

Mike will visit us tomorrow.	They went to Japan last summer.
Janet will bring salmon tomorrow.	She will start a new job next month.
Ken never reads his mails.	They can come for dinner tonight.
Jerry went to the movies last night.	I received the pictures yesterday.
He used to play the drums in this band last year.	Jane traveled with her friend to India.