Mateusz Jekiel

331039

The influence of musical hearing on foreign language pronunciation in Polish advanced learners of English

Wpływ słuchu muzycznego na wymowę języka obcego u polskich uczniów z zaawansowaną znajomością języka angielskiego

Rozprawa doktorska napisana
na Wydziale Anglistyki
Uniwersytetu im. Adama Mickiewicza w Poznaniu
pod kierunkiem dr hab. Piotra Gąsiorowskiego, prof. UAM

Poznań, 2023

I dedicate this work to my mother, Aleksandra,
for inspiring me to take this journey.

Dedykuję tę pracę mojej mamie, Aleksandrze,
za zainspirowanie mnie do tej podróży.

ACKNOWLEDGEMENTS

First of all, I would like to express my sincere gratitude to my esteemed supervisor, dr hab. Piotr Gąsiorowski, prof. UAM, for his constant support, insightful comments and valuable suggestions. Thank you for giving me the freedom in my scientific undertakings, allowing me to follow my research interests and believing in my dreams. My gratitude extends to the Faculty of English for the opportunity to undertake my PhD studies at the Adam Mickiewicz University in Poznań, Poland. I would also like to thank prof. Katarzyna Dziubalska-Kołaczyk for her treasured support in my academic endeavours.

I would like to offer my special thanks to Kamil Malarski, my dear friend and colleague, who has been supporting me since the very beginning. Who would have thought that two classmates from primary school would end up working together in the same phonetics laboratory, recording participants and performing acoustic analyses. Your dedication and kindness always inspired me to become a better researcher and a better person.

I would also like to thank Kacper Łodzikowski for his support in my research endeavours. Your willingness to share your knowledge and expertise always motivated me to be open and transparent in my research. Thank you for all your teachings and help.

I am deeply grateful to my family, especially my parents, Aleksandra and Jacek Jekiel, and my grandmother Krystyna Krajniak, for their love and encouragement.

Finally, words cannot express my gratitude to Halina Lewandowska for her unconditional faith and support in my pursuits. I thank you from the bottom of my heart.

Table of contents

List of Research articles

The present PhD thesis comprises three thematically related Research articles:

Research article 1 (Jekiel and Malarski 2021)

Jekiel, Mateusz and Kamil Malarski. 2021. "Musical hearing and musical experience in second language English vowel acquisition", *Journal of Speech, Language, and Hearing Research* 64, 5: 1666-1682. https://doi.org/10.1044/2021_JSLHR-19-00253

Research article 2 (Jekiel 2022)

Jekiel, Mateusz. 2022. "L2 rhythm production and musical rhythm perception in advanced learners of English", *Poznan Studies in Contemporary Linguistics* 58, 2: 315-340. https://doi.org/10.1515/psicl-2022-0016

Research article 3 (Jekiel and Malarski 2023)

Jekiel, Mateusz and Kamil Malarski. 2023. "Musical hearing and the acquisition of foreign-language intonation", *Studies in Second Language Learning and Teaching* 13, 1: 151-178. https://doi.org/10.14746/ssllt.23166

Funding

The research conducted for the present PhD thesis and its three published Research articles (Jekiel and Malarski 2021; Jekiel 2022; Jekiel and Malarski 2023) was funded by the National Science Centre in Poland under a research grant entitled "Musical hearing in the acquisition of EFL pronunciation" (Preludium, 2014/15/N/HS2/03865). The principal investigator of this research grant was the author of the present PhD thesis and the author of Research article 2 (Jekiel 2022), as well as the first author of Research article 1 (Jekiel and Malarski 2021) and Research article 3 (Jekiel and Malarski 2023). The co-investigator was the second author of Research article 1 (Jekiel and Malarski 2021) and Research article 3 (Jekiel and Malarski 2023). The recipient of the grant was Adam Mickiewicz University in Poznań, Poland.

List of abbreviations

CVC	consonant-vowel-consonant
EFL	English as a foreign language
F_0	fundamental frequency
F_1	first vowel formant
F_2	second vowel formant
GA	General American
GB	General British
IP	intonational phrase
L1	first language
L2	second language

Part 1: Introduction

The relationship between language and music has been discussed over many years, from Rousseau's (1781) hypothesis that speech evolved from singing and Darwin's (1871) argument that musical sounds are the foundation for the development of language. Contemporary research in the field of language evolution suggests that the development of musical abilities among humans was similar to their language capacity (Brown 2000), or even that language originated from music (Mithen 2005). Studies exploring the similarities in language and music point to a number of commonalities between the two domains: both exist in all cultures (Nettl 2000); both share the same neural resources for processing (Brown et al. 2006); both require memory and sensorimotor coordination (see Besson et al. 2011 for a review); both could be described as complex and organised auditory signals containing tone, melody, and rhythm, as well as measurable frequency, duration, and timbre (see Fadiga et al. 2009 for a review). Despite these shared characteristics, Jackendoff (2009) underlined their correlational nature and highlighted the growing need to analyse other capacities to produce an objective foundation for these comparisons. One such field is the study of the relationship between musical hearing and foreign language pronunciation.

Language researchers have investigated the influence of musical abilities on various elements of non-native pronunciation (e.g. Milovanov et al. 2010; Zybert and Stępień 2009), including vowels (e.g. Intartaglia et al. 2017; Kempe et al. 2015), rhythm (e.g. Gralińska-Brawata and Rybińska 2017; Magne et al. 2016), and intonation (e.g. Patel et al. 2005; Zatorre and Baum 2012). However, the sheer variability of methods and previous research limitations still leaves room for further investigation, as former studies frequently relied on self-reported language proficiency (e.g. Roncaglia-Denissen et al. 2016), impressionistic assessment (e.g. Milovanov et al. 2010), speech shadowing (e.g. Pastuszek-

Lipińska 2008), or formal music education (e.g. Marques et al. 2007). Although the notion of musical aptitude's impact on foreign language learning is present in current research, there has been no longitudinal study that would quantitatively assess the exact influence of particular aspects of musical hearing on individual substrates of non-native pronunciation in a classroom setting.

The objective of this PhD project is to investigate if there is a significant relationship between specific aspects of musical hearing and the acquisition of selected features of foreign language pronunciation in a formal learning environment. This research question is addressed in three thematically connected Research articles. Each Research article inspected an aspect of non-native speech of 50 Polish undergraduate students of English in a longitudinal study spanning two academic semesters with the use of acoustic analysis and quantitative measurements, along with musical hearing tests and a questionnaire on former musical experience.

This first part of the present PhD thesis introduces the theoretical background, followed by a summary of the primary research objectives and detailed descriptions of Research article 1 (Jekiel and Malarski 2021), Research article 2 (Jekiel 2022), and Research article 3 (Jekiel and Malarski 2023). Each description includes theoretical background, research objectives, results and their interpretation, study limitations, and conclusions.

1.1. Theoretical background

The following section covers the concepts that are crucial to the present PhD thesis, supported by relevant research on the influence of musical hearing and musical experience on foreign language pronunciation.

1.1.1. The sound of music and speech

Over the years, interdisciplinary research in language and music has grown into a vast field, comprising works on language and music evolution (e.g. Brown 2000; Fitch 2010), music and neurolinguistics (e.g. Brown et al. 2006; Patel 2008), as well as the role of musical skills in foreign language aptitude (see Turker and Reiterer 2021 for a review). Several studies demonstrated the interplay between music and speech (see Slevc 2012 for a review) and the influence of musical hearing on language acquisition (see Brandt et al. 2012 for a review). Indeed, the sounds of music and speech share the same acoustic properties conveying information, including pitch (i.e. the perceptual equivalent of fundamental frequency), timing (i.e. the temporal features of sounds), and timbre (i.e. the quality of sounds) (Kraus and Chandrasekaran 2010). These similarities have led researchers to investigate the relationship between musical hearing and speech perception.

Many studies have confirmed the connection between musical training and pitch perception in L1 (e.g. Besson et al. 2007; Schön et al. 2004) and L2 (e.g. Deguchi et al. 2012; Marques et al. 2007), as well as the connection between musical training and phonological awareness (e.g. Dege and Schwarzer 2011; Flaugnacco et al. 2015), including the processing of foreign-speech sounds (e.g. Intartaglia et al. 2017; Götz et al. 2023). Similarly, research in musical aptitude has revealed a strong link between musical hearing and pitch perception (e.g. Deguchi et al. 2012), phonological awareness (e.g. Culp 2017), as well as speech processing in L1 (e.g. Strait et al. 2012) and L2 (e.g. Kempe et al. 2015). Additional studies in music and speech processing have also shown parallels across other aspects, such as rhythm perception (e.g. Magne et al. 2016), stress perception (Choi 2022), and detection of small differences between speech sounds (Parbery-Clark et al. 2012). Overall, previous research has demonstrated the complex interactions between musical aptitude, musical training, and various aspects of speech perception.

1.1.2. Musical hearing and musical experience

For the present PhD thesis, it is important to note the difference between two terms used across previous studies, namely musical training and musical aptitude. The former, labelled also as musical practice or musical experience, can be regarded as the skill and knowledge acquired by performing music through singing or playing a musical instrument, and is often related to formal music education. The latter, also termed musical ability (Bentley 1969) or musical talent (Seashore et al. 1960), is the capacity to structure acoustic information (Karma 2007); it is linked to musical hearing, i.e. the perceptual sensitivity to acoustic phenomena in music, such as pitch or rhythm, and is separate from musical knowledge (Wallentin et al. 2010). Recent research suggests that musical hearing might be a better predictor of successful acquisition of language skills than formal music education (Choi 2022; Swaminathan and Schellenberg 2020), since defining a musician could be problematic due to the idiosyncratic differences in musical training (Law and Zentner 2012) or innate talent (Nardo and Reiterer 2009), as well as expertise in a particular musical instrument (Gottfried 2007) or singing (Christiner and Reiterer 2016). However, musical experience should not be merely discarded from future analyses, but rather included as one of many factors that constitute the complex relationship between musical skills and language skills.

The present PhD thesis adopts the terms musical hearing and musical experience, since these terms were used across all three Research articles (Jekiel and Malarski 2021; Jekiel 2022; Jekiel and Malarski 2023). First, the decision to use the term musical hearing as opposed to musical aptitude or musical talent stems from its focus on the perceptual aspect, rather than denoting a more general skill or gift, which could also be related to singing or playing a musical instrument. Second, the use of the term musical experience instead of musical training or musical practice is motivated by its inclusion of both formal music education as well as amateur music background, since both could be potential factors in the acquisition of foreign language pronunciation.

In the study, musical hearing was assessed by means of three musical hearing tests designed by Mandell (2009): the Adaptive Pitch Test measuring pitch perception abilities, the Rhythm Test measuring rhythmic memory, and the Tonedeaf Test measuring melodic memory. During data collection, these musical hearing tests were easily accessible online and relatively manageable, especially when compared to previous musical hearing tests (e.g. Gordon 1989; Seashore et al. 1960), which have been criticised for their difficulty and inaccessibility (see Law and Zentner 2012 for a review). Moreover, the tests designed by Mandell have recently been used in similar studies investigating the role of musical hearing in non-native speech perception (e.g. Lorenzen 2019; Ning 2020). Aside from the musical hearing tests, musical experience was assessed by means of a questionnaire, in which participants reported their prior music education and musical practice, as well as the years of that musical experience. The assessment of pre-existing musical hearing skills and former musical experience was assumed to produce a more comprehensive image of what is otherwise broadly defined as musicality.

1.1.3. Musical hearing and musical experience in foreign language pronunciation

In an increasingly multilingual world, communicating in a foreign language has become one of the most essential skills in professional development. In particular, the role of pronunciation in L2 speech is not only limited to comprehension but is also a determining factor in the perception of L2 speakers' competence, education, and intelligence (Dewaele and McCloskey 2015). At the same time, achieving native-like L2 pronunciation is considered one of the most challenging aspects of foreign language learning, especially among adult L2 learners, whose L2 proficiency is subject to variation due to individual differences in their L2 learning aptitude (Doughty 2019). Therefore, it is necessary to determine which factors can help in successful acquisition of native-like L2 pronunciation. Previous studies indicate a wide range of factors related to foreign language pronunciation (see Suzukida 2021 for a review), such as age of acquisition (Saito 2015), language talent (Dogil and Reiterer 2009), and motivation (Smit 2002), as well as musical hearing (Delogu and Zheng 2020; Slevc and Miyake 2006) and musical experience (Borodkin et al. 2022; Chobert and Besson 2013), which are the focal points of the present PhD thesis.

According to the Speech Learning Model (Flege 1995; Flege and Bohn 2021), the acquisition of L2 pronunciation begins on a perceptual level. When L2 learners start to discriminate the acoustic dimensions of L2 sounds (e.g. pitch, length, quality) from their L1 counterparts, new phonetic categories take form in their minds. Since accurate production of non-native sounds is correlated with their perception, the improved performance in perception should also lead to the improvement of foreign-language pronunciation. In general, L2 pronunciation can be described as the ability to accurately produce the segmental (i.e. vowels and consonants) and suprasegmental (i.e. rhythm and intonation) elements of foreign pronunciation. Since foreign language pronunciation among adult learners is often foreign-accented due to their L1 influence (Flege and Bohn 2021), it has been claimed that L2 pronunciation teaching and learning should be focused on intelligibility, rather than nativeness (Levis 2018). However, achieving native-like pronunciation is still desired among many foreign language learners (e.g. Burri 2023; Richter 2021), including Polish students of English (e.g. Nowacka 2022; Waniek-Klimczak et al. 2015). While some studies have shown that foreign language learners can improve in their comprehensibility (Derwing and Munro 2013), achieving native-like L2 pronunciation can be difficult to attain (Derwing et al. 2014). Therefore, reviewing factors affecting non-native speech should be of paramount importance in second language research.

The positive transfer of musical hearing skills to foreign language pronunciation is in agreement with former studies investigating the perception-production link in L2 acquisition (see Sakai and Moorman 2018 for a review). Extrapolating from the Speech Learning Model (Flege 1995), it is possible to predict that musical aptitude can lead to more accurate perception of L2 speech, which in turn can lead to more accurate L2 speech production. Furthermore, since the revised Speech Learning Model (Flege and Bohn 2021) assumes that the relationship between perception and production is bidirectional, it is also possible to suggest that musical hearing and foreign language pronunciation may interchangeably influence each other (see Borodkin et al. 2022 for a similar assumption).

Previous studies investigating the link between language and music processing focused predominantly on the perceptual aspects (e.g. Kempe et al. 2015), specifically pitch perception (e.g. Zatorre and Baum 2012) or rhythm perception (e.g. Magne et al. 2016). At the same time, only a limited number of studies reported that musical hearing and musical experience can improve L2 pronunciation. Slevc and Miyake (2006) observed a positive correlation between musical hearing and L2 production in Japanese adult learners of English. Milovanov et al. (2010) reported a link between musical aptitude and pronunciation skills in Finnish adult learners of English in the production of selected English phonemes. In research involving Polish learners of English, Zybert and Stępień (2009) reported a correlation between musical aptitude and the production of selected elements of foreign language pronunciation in a small sample. Gralińska-Brawata and Rybińska (2017) suggested a possible link between musical abilities and correct L2 word stress production, declaring the need for a larger group study. Finally, Kaszycka (2021) demonstrated a moderate correlation between perceived pronunciation proficiency and musical ability.

The above-mentioned studies not only provided novel insights into the complex relationship between musical hearing skills and non-native speech, but also opened new perspectives for future research. Many former studies focused only on participants with and without formal music education, while recent research suggests that musical hearing skills may be more conducive to L2 acquisition than musical training (see Lund 2023 for a review). Moreover, pronunciation data has often been obtained via speech shadowing tasks and frequently assessed impressionistically, which could be limiting and misrepresent learners' actual pronunciation skills (Dufour and Nguyen 2013). Despite the recognised impact of musical aptitude on foreign language acquisition, there has still not been a longitudinal study that would quantitatively assess the exact influence of particular aspects of musical hearing on individual substrates of non-native pronunciation in a formal learning environment.

1.2. Research objectives

Encouraged by previous research on the role of musical aptitude in L2 speech acquisition, this PhD project investigates the relationship between specific aspects of musical hearing and specific features of foreign language pronunciation. Each Research article focused on a different element of pronunciation in order to provide a better understanding of the connections between musical hearing and L2 speech. Specifically, the objective of the Research articles was to explore the links between pitch perception, rhythmic memory, and melodic memory on the acquisition of L2 vowels, rhythm, and intonation in a longitudinal study among Polish undergraduate students of English. To this end, 50 advanced learners of English aged 19-21 received accent training in the form of a two-semester practical course in English pronunciation supplemented with a theoretical course in English phonetics and phonology. Participants were recorded using SpeechRecorder (Draxler and Jänsch 2004) before and after training, completed a set of musical hearing tests (Mandell 2009) and a questionnaire on prior music education and musical experience. Research article 1 (Jekiel and Malarski 2021) investigated whether pre-existing musical hearing skills and prior musical experience can predict native-like pronunciation of L2 vowels in monosyllabic words. Research article 2 (Jekiel 2022) focused on the role of musical hearing and musical experience in the acquisition of acoustic correlates of L2 rhythm using a reading passage. Research article 3 (Jekiel and Malarski 2023) studied the relationship between musical aptitude and accurate production of L2 intonation patterns in read dialogues. Individual research questions are presented in detail in the respective Research articles.

1.3. Research article 1 (Jekiel and Malarski 2021): Musical hearing and musical experience in second language English vowel acquisition

There are several linguistic and extralinguistic factors contributing to the successful acquisition of L2 pronunciation. While the former focus on potential similarities and differences between the phonetic inventories of L1 and L2, the latter encompass such aspects as the age of onset, learner motivation, learning strategies, and working memory. According to Chobert and Besson (2013), musical hearing is among such extralinguistic factors, and can positively influence both speech perception and production. Unlike formal music education or practical musical experience, musical hearing can be defined as an untaught natural musical ability (Dolman and Spring 2014) and an innate sensitivity to such aspects of music as pitch or timing, which are also characteristic of vowels (Kraus et al. 2009). According to the studies on the effect of musical hearing on L1 and L2 pronunciation, musical aptitude can positively impact L1 phonological processing in children (Anvari et al. 2002), as well as improve the perception and production of L2 minimal pairs in adult learners (Slevc and Miyake 2006), and native-like pronunciation of challenging L2 English vowels and consonants (Milovanov et al. 2010). In particular, vowels can be regarded as very music-like phonemes, as Fenk-Oczlon (2017) describes their role in generating sonority across syllables in both speech and singing. Like musical notes, vowels have duration, intensity, pitch, and timbre (Lidji et al. 2010), and their musical quality plays a vital function in early language acquisition (Masataka 2007). The relationship between musical hearing and L2 vowel acquisition was examined in a series of studies by Milovanov et al. (2008; 2009; 2010), who propose that the perception and production of L2 vowels are interconnected with musical hearing. Despite the growing number of studies on the connection between musical hearing and L2 acquisition, only a limited number of studies have examined the effect of musical skills on successful L2 pronunciation across Polish learners of English (Gralińska-Brawata and Rybińska 2017; Pastuszek-Lipińska 2008; Zybert and Stępień 2009), focusing primarily on musicians with formal musical training and relying on imitation and speech shadowing.

Research article 1 (Jekiel and Malarski 2021) aimed to fill this gap and explore the influence of musical hearing, music education, and musical experience on the successful acquisition of L2 English vowels in a longitudinal study. The participants were 50 Polish undergraduate students of English aged between 19 and 21, they were all advanced learners of English as confirmed by their LexTALE scores (Lemhöfer and Broersma 2012) and had no prior formal L2 pronunciation instruction. All participants received accent training in the form of a two-semester practical course in English pronunciation, which aimed at developing a consistent and native-like GB accent by practising segmental and suprasegmental aspects of L2 pronunciation. The practical course was also supplemented with a theoretical course in English phonetics and phonology. The participants were recorded before and after training, reading aloud a series of short monosyllabic words including GB monophthongs in CVC contexts. The recordings were pre-examined in Praat (Boersma and Weenink 2023) and automatically segmented in DARLA (Reddy and Stanford 2015), using the Montreal Forced Aligner (McAuliffe et al. 2017), FAVE-Extract (Rosenfelder et al. 2014), and the Vowels R package (Kendall and Thomas 2010). The extracted vowel formants were then compared with the target GB vowel formants produced by the learners' pronunciation teachers, whose formant frequencies were similar to model GB scores found in Cruttenden (2014). Before training, all participants completed three musical hearing tests (Mandell 2009) assessing pitch perception, rhythmic memory, and melodic memory, as well as a questionnaire on former music education and musical experience. Afterwards, the mean Euclidean distance between participants' vowel formants and their pronunciation teachers' GB vowel formants were measured and juxtaposed with the musical hearing test results and the questionnaire responses. Taking previous research into account, the current study expected to observe more native-like production of L2 vowels after training, particularly among participants with better musical hearing or musical experience. The study aimed to answer the following research questions: (1) Is there an observable improvement in the production of L2 vowels after training? (2) Do participants with better musical hearing produce more native-like L2 vowels? (3) Can musical experience be used to predict success in learning to produce L2 vowels?

The results showed a significant difference in the production of L2 vowels before and after training, as participants' mean L2 vowel formants were closer to the model GB vowel formants of their pronunciation teachers after the two-semester accent training. However, not all GB vowels were acquired with the same degree of success; while the participants' post-training formant values of front monophthongs were the closest to the model, the formant values of back monophthongs were the most distanced. Secondly, two linear mixed-effects regression models were built to explain the closeness of participants' L2 vowel formants to the pronunciation teachers' GB vowel formants before and after training.

In both models, gender, LexTALE scores, musical hearing test scores, music education, and years of musical experience were fixed effects, while speaker and vowel were random effects. The results for the first model showed that the rhythmic memory test scores were good estimates for the proximity of participants' L2 vowel formants to the teachers' formant values before training, i.e. students with higher rhythmic memory test scores produced more native-like GB vowels before the accent training. The strongest observable estimate in the second model was the similarity to teachers' GB vowel formants before training, i.e. participants with more native-like L2 vowels prior to the accent training were more successful after the two-semester accent training. No effect was found between pitch perception test scores or melodic memory test scores and the closeness to the model GB vowels before or after training. Finally, a significant result was found between the proximity to the model pronunciation after training and musical experience, i.e. participants who spent more years singing or playing a musical instrument were able to produce more native-like GB vowels after the accent training. No effect was found between the closeness to the model formant values before or after training and music education, albeit only four out of fifty participants reported attending music school.

The results of the study suggest not only that L2 pronunciation is teachable and learnable in an academic setting, but also that rhythmic memory and musical experience can be significant predictors of successful L2 vowel acquisition. While rhythm is usually associated with speech prosody, studies pointed to the potential link between rhythm perception and the temporal features of the English tense-lax contrast (Schwartz 2010), as well as its importance in language acquisition (see Langus et al. 2017 for a review) and L2 pronunciation instruction (Llanes-Coromina et al. 2018). The results are also in line with studies where rhythm perception was found to be related to the production of fewer L2 pronunciation errors (Milovanov et al. 2010), as well as studies on the use of musical rhythm in teaching durational contrasts (Wang et al. 2016). Since vowel reduction and durational contrasts are commonly found as problematic for Polish learners of English, the results of this study suggest that rhythmic memory can help alleviate those issues and help EFL learners achieve more native-like pronunciation. Finally, the connection between the mean vowel formants before and after training reveal that participants' former proficiency in L2 pronunciation was a strong indicator for their improvement after training. This finding suggests that while musical hearing and musical experience can be related to successful acquisition of L2 vowels, individuals' pre-existing L2 pronunciation skills can be a better predictor of achieving native-like GB pronunciation during a two-semester accent training in a formal learning environment.

The limitations of the study are threefold. First, the specific aspects of the acquisition of L2 vowels by Polish learners of English may not be easily transferred to other languages; while Polish EFL learners frequently struggle with the TRAP vowel (Weckwerth 2011) or the KIT–FLEECE vowel contrast (Rojczyk and Porzuczek 2012), EFL learners with different L1s may find the acquisition of other L2 vowels more difficult. Second, while the study focused on comparing formant values, i.e. vowel quality, it would be of great interest to investigate the relationship between rhythmic memory and vowel quantity, as vowel duration and vowel reduction are important aspects of English phonology and new concepts in the acquisition of L2 pronunciation by Polish learners of English (this is further investigated in the next Research article in this series). Finally, the participants of the study were Polish students of English, and the research was conducted in an academic setting, with no control for music education or musical experience. Since the starting age of musical training can be related to successful language acquisition (Brandt et al. 2012), the ability to control for this variable in future research may yield more comprehensive results.

Research article 1 (Jekiel and Malarski 2021) reported a longitudinal study investigating the role of pre-existing musical hearing skills and prior musical experience in the acquisition of L2 vowels by Polish EFL learners over a two-semester accent training. The results revealed that rhythmic memory and musical experience can be significant factors in the production of more native-like L2 vowels before and after training, respectively. While the mean Euclidean distance between the participants' vowel formant values and the model GB vowel formant values generally decreased after training, the improvement was more significant among participants who had some experience in singing or playing a musical instrument, suggesting that musical experience may be an asset in the acquisition of L2 vowels. Finally, participants' initial proficiency in L2 pronunciation was the strongest predictor of acquiring native-like GB vowels during the accent training, suggesting that musical hearing or musical experience may have a secondary role when learning a foreign language accent in a formal learning environment. Despite its limitations, the study hopefully adds to available literature on the role of musical hearing and musical experience in L2 speech.

1.4. Research article 2 (Jekiel 2022): L2 rhythm production and musical rhythm perception in advanced learners of English

Recent research regarding the relationship between language rhythm and musical rhythm still remains inconclusive. Both domains use pitch and duration for labelling group boundaries within their complex and hierarchical rhythmic structures (Patel 2008) and share neurocognitive resources for their processing (Magne et al. 2016), implying that musical rhythm could be an offshoot of language rhythm. These commonalities allowed to investigate the role of rhythm perception in language acquisition (see Langus et al. 2017 for a review) and phonological awareness (Moritz et al. 2013). Studies extending their scope to L2 speech argue that musical ability can positively influence L2 rhythm perception (Choi 2022), L2 comprehensibility (Llanes-Coromina et al. 2018), L2 pronunciation proficiency (Kaszycka 2021) and L2 prosody (Cason et al. 2020). However, studies on musical hearing and L2 rhythm in Polish EFL learners are still scarce (see Gralińska-Brawata and Rybińska 2017). Also, no study has yet implemented the use of rhythm metrics to explore the relationship between musical hearing and L2 rhythm production, relying on impressionistic judgements of speech instead.

The aim of Research article 2 (Jekiel 2022) was to investigate the link between musical hearing and the acquisition of L2 rhythm by Polish EFL learners using quantitative rhythm metrics and musical aptitude measurements in a longitudinal study. To this aim, 50 Polish advanced learners of English were recorded reading *Please Call Stella* (elicitation paragraph used in phonetic studies, see Weinberger 2015) before and after training. The participants also completed two musical hearing tests assessing rhythmic memory and melodic memory (Mandell 2009), followed by a questionnaire on musical experience. All recordings were pre-examined in Praat (Boersma and Weenink 2023) and automatically segmented in DARLA (Reddy and Stanford 2015), using the Montreal Forced Aligner (McAuliffe et al. 2017), FAVE-Extract (Rosenfelder et al. 2014), and the Vowels R package (Kendall and Thomas 2010). The rhythm metrics (Vdev, Cdev, VarcoV, VarcoC, nPVI-V, nPVI-C, CCI-V and CCI-C) were calculated using Correlatore (Mairano and Romano 2010) and the obtained scores before and after training were then compared with the pronunciation teachers' rhythm metric scores. Finally, the proximity between the participants' rhythm metric scores and the model rhythm metric scores was correlated with the musical hearing test scores and years of musical experience. Due to the established differences in the vocalic intervals for English and Polish (Grabe and Low 2002; Mairano and Romano 2011), the present study expected to observe a difference in the rhythm metric scores for Polish EFL learners before and after the accent training, especially among participants with higher musical hearing test scores. The study aimed to answer the following research questions: (1) Can rhythm metrics be used to observe L2 rhythm acquisition in a longitudinal study? (2) Can rhythm metrics be used to distinguish intermediate EFL learners from advanced EFL learners in their production of L2 rhythm? (3) Is there a relationship between musical hearing or musical experience and L2 rhythm acquisition in a formal learning environment?

The results of the study revealed a significant difference across all rhythm metric scores between vocalic intervals (Vdev, VarcoV, nPVI-V, CCI-V) before and after training. i.e. participants produced more native-like L2 vowel duration contrasts and L2 vowel reduction after the accent training, leading to higher vocalic variation and, accordingly, higher vocalic rhythm metric scores. At the same time, there was no significant difference found between most consonantal rhythm metric scores (Cdev, VarcoC, nPVI-C) before and after training. The significant result for CCI-C indicates that participants produced more native-like L2 consonant clusters after training, leading to lower consonantal variation. This is also in agreement with previous research using this rhythm metric (Mairano and Romano 2011). Finally, no significant correlation was found between the rhythm metric scores and musical hearing test scores or musical experience, aside from a weak positive correlation between the rhythmic memory test scores and post-training VarcoV and nPVI-V scores.

The observed progress in the production of higher vocalic variation across all rhythm metric scores suggests that a two-semester accent training for Polish EFL learners can result in more native-like L2 production of vowel duration contrasts and L2 vowel reduction, i.e. aspects of pronunciation directly connected to perceived speech rhythm. These results confirm the application of rhythm metrics in the study of L2 rhythm acquisition in a formal learning environment. Next, the lack of a significant link between the rhythm metric scores and musical hearing test scores or musical experience suggests that musical aptitude or training may not be a strong factor when the acquisition of L2 rhythm takes place during an intensive two-semester accent training at an academic level. In other words, the observed progress in L2 rhythm production may result from the accent training alone, allowing all participants to gain from explicit instruction, regardless of pre-existing musical hearing skills or prior musical experience. Alternatively, it is possible that musical hearing can be related to other aspects of language rhythm not included in rhythm metrics, such as fundamental frequency or intensity (Cumming 2011). Indeed, language rhythm is a complex phenomenon, which is difficult to measure and shaped by various linguistic and extra-linguistic factors.

There are three limitations to this study that should be noted. First, rhythm metrics rely on durational contrasts and do not include other factors affecting perceived speech rhythm, such as pitch (Pickering and Wiltshire 2000), sonority (Galves et al. 2002) or loudness (Fuchs 2014). Including these aspects should provide a more comprehensive outline of language rhythm. In addition, Arvaniti (2012) and Gut (2012) list a number of study design factors that can affect L2 rhythm metrics, including the material, speaking style, speech rate, and language proficiency, which make research reproducibility challenging. Secondly, an important factor in L2 acquisition among adult EFL learners is motivation, which can be a strong predictor of achieving native-like L2 pronunciation (Smit 2002). Although the participants of this study were a homogeneous group of adult advanced learners of English in an academic setting, it is possible that highly motivated participants might have spent more time practising on their own, compensating for their lack of pre-existing musical hearing skills or prior musical experience. Thus, it would be vital to control participants' motivation and self-study in future longitudinal studies on L2 rhythm. Finally, incorporating tests assessing music production (Wallentin et al. 2010) and tests designed to evaluate rhythmic skills in greater detail (Bella et al. 2017) should produce a more complete picture of the relationship between musical aptitude and L2 rhythm acquisition.

The findings revealed in Research article 2 (Jekiel 2022) suggest that rhythm metrics can be used in a longitudinal study to detect progress in the acquisition of L2 rhythm, adding to previous research in this field (Gralińska-Brawata 2014). The results also confirm that L2 rhythm is teachable and learnable in a formal learning environment during a two-semester accent training. At the same time, no significant relationship between L2 rhythm metric scores and musical hearing or musical experience was found. However, it is possible that these aspects of musicality may be connected to other attributes of speech rhythm or other elements of prosody, such as intonation (investigated further in the next Research article in this series). Hopefully, this study can help direct future quantitative and longitudinal research in L2 rhythm and the role of musical hearing in L2 pronunciation.

1.5. Research article 3 (Jekiel and Malarski 2023): Musical hearing and the acquisition of foreign-language intonation

In the wide field of research in language and music, intonation in speech and melody in music are considered the two most closely aligned facets. First, both aspects rely on the same acoustic parameters and spectral characteristics (Schön et al. 2004). Second, intonation and melody in speech and music are processed by the same "coarse-grained" system functioning in the brain (Zatorre and Baum 2012). These commonalities lead researchers to investigate the potential influence of pitch perception in speech and music on language acquisition. Perception-based research confirmed that musical pitch perception is related to encoding of speech contours and tone discrimination in L2 speech (Wong et al. 2007; Marie et al. 2010), while studies comparing musicians to non-musicians showed that musically trained foreign language learners have better reaction times for auditory processing of speech tones (Ott and Jäncke 2013) and more native-like listening skills (Intartaglia et al. 2017). However, research on the relationship between musical pitch perception and the production of L2 intonation is still sparse. In a study by Pastuszek-Lipińska (2008), musically trained Polish EFL learners produced fewer pronunciation errors in a speech shadowing task, albeit received similar scores for their L2 intonation to non-musicians. Zybert and Stępień (2009) reported that Polish EFL learners with superior musical hearing test scores also received higher scores in speech shadowing tasks assessing L2 prosody, including word stress and intonation. Both studies relied on impressionistic judgement of speech and compared EFL learners with and without formal music education.

The primary goal of Research article 3 (Jekiel and Malarski 2023) was to investigate the relationship between musical hearing and the production of L2 intonation by Polish advanced learners of English in a longitudinal study using quantitative measures. The secondary goal of the study was to determine whether L2 intonation is learnable and teachable in a formal learning environment. To this end, 50 Polish advanced learners of English were recorded before and after training. The stimulus was a series of short dialogues adapted from Wells (2014) and designed to elicit different English intonation patterns. Next, the participants completed two musical hearing tests assessing pitch perception and melodic memory (Mandell 2009), as well as a questionnaire on musical experience. Participants' fundamental frequency contours were visually analysed using ToBI guidelines (Beckman and Elam 1997) and compared with their pronunciation teachers' intonation patterns. The percentages of correctly produced intonation patterns were compared with the musical hearing test scores and questionnaire responses using linear multiple regression models and two-way ANOVAs. Following previous research in this field, this study expected to observe a potential relationship between accurate production of L2 intonation and pre-existing musical hearing skills. The study tried to answer the following research questions: (1) Is there an observable difference in the production of L2 intonation patterns after a two-semester accent training? (2) Is there a relationship between musical hearing and accurate production of L2 intonation? (3) Is musical experience connected to the acquisition of L2 intonation?

The results of the study revealed apparent improvement in L2 intonation after training, i.e. participants produced more native-like GB intonation patterns after the two-semester accent training, similarly to their pronunciation teachers. However, not all intonation patterns were acquired with the same degree of success: the fall-rise and the rise-fall patterns were the most difficult to acquire, while in wh-questions and tag questions the correct falling pattern was produced more often than a rise after training. Next, linear multiple regression results revealed that musical pitch perception was a significant predictor of accurate production of L2 intonation patterns after training, while melodic memory test results had no effect on post-training intonation scores. No significant correlation was found between participants' musical experience and accurate production of L2 intonation. Finally, two-way ANOVAs were performed to analyse the effect of musical hearing and musical experience on the production of L2 intonation patterns before and after training. No significant results were reported, but participants with good musical hearing test scores produced more native-like intonation patterns after the accent training.

The findings show that L2 intonation is both learnable and teachable in a formal learning environment, as participants significantly improved after the two-semester accent training by replicating their pronunciation teachers' model intonation patterns. Second, the significant result for pitch perception in the linear mixed regression suggests that having a good musical ear can be beneficial in the process of learning foreign language intonation. Indeed, accurate pitch change recognition can be an important skill when practising intonation in the EFL classroom, where such terms as "rises" and "falls" can be frequently used (Zybert and Stępień 2009). Finally, contrary to former studies on the role of musical training in the acquisition of L2 intonation (Pastuszek-Lipińska 2008), music education and musical experience were not observed as significant factors in this study, implying that musical background may not be as crucial when learning L2 intonation in a formal learning environment. Alternatively, the result may suggest that intensive accent training allowed all learners to progress, regardless of their former musical experience.

The following limitations to this study should be acknowledged. First, as contemporary EFL learners are subject to a great variety of audio-visual content in native-spoken English, it is difficult to determine a single factor affecting L2 pronunciation skills. While this study focused primarily on the fundamental frequency present in intonation patterns, speech prosody also includes pitch register, pitch span, and speech rhythm. Thus, it would be of great interest to include other prosodic features in future research investigating the role of musical hearing in L2 pronunciation. Second, as participants were Polish EFL students studying at the same university, there was no control for their prior music education and musical experience, similarly to Research article 1 (Jekiel and Malarski 2021) and Research article 2 (Jekiel 2022). Finally, performing additional tests assessing musical training, either in singing or playing a musical instrument, may be a valuable addition to this and the above-mentioned studies.

To summarise, Research article 3 (Jekiel and Malarski 2023) reported a longitudinal study exploring the connection between musical hearing and L2 intonation after a two-semester accent training for Polish advanced learners of English using visual analysis of the fundamental frequency, musical hearing tests, and a questionnaire on musical experience. The results of the study showed that higher scores in the pitch perception test were related to post-training accuracy in the production of native-like L2 intonation patterns, similar to the participants' pronunciation teachers. However, students with lower scores in musical hearing tests and no musical experience also improved, suggesting that acquiring L2 intonation in a formal learning environment is possible for all learners of English, regardless of their musical hearing or musical experience. More research is still required to assess the relationship between musical aptitude and other factors affecting speech prosody, as well as more control for prior music education and musical training.

Part 2: General Discussion

The second part of the present PhD thesis provides a general discussion of the findings from Research article 1 (Jekiel and Malarski 2021), Research article 2 (Jekiel 2022), and Research article 3 (Jekiel and Malarski 2023). The obtained data and results concerning specific elements of foreign language pronunciation, musical hearing, and musical experience are evaluated across studies. In addition, research limitations and future research directions are also taken into consideration. All individual results are discussed with reference to the research questions as formulated in each respective Research article.

2.1. The acquisition of foreign language pronunciation

In all three studies of this PhD project participants' L2 speech was recorded before and after a two-semester accent training focusing on GB pronunciation. All three studies recruited the same pool of participants, i.e. 50 Polish undergraduate students of English. The dataset in Research article 1 (Jekiel and Malarski 2021) comprised GB monophthongs in a CVC context produced by each participant in order to perform acoustic analyses of L2 vowel formants. In Research article 2 (Jekiel 2022), each participant read *Please Call Stella* (elicitation paragraph used in phonetic studies, see Weinberger 2015) in order to elicit L2 rhythm and calculate rhythm metrics. In Research article 3 (Jekiel and Malarski 2023), each participant read short dialogues adapted from Wells (2014) to obtain L2 intonation patterns. Significant changes after training were reported across all three studies. Specifically, the participants' post-training vowel formants, rhythm metric scores, and intonation patterns, were more similar to those produced by their pronunciation teachers. These findings are in line with the research questions found in each Research article and confirm that L2 vowel quality, L2 rhythm, and L2 intonation are teachable and learnable in a formal academic environment during a two-semester accent training. The following sections discuss in detail the acquisition of each aspect of pronunciation during the accent training.

2.1.1. The acquisition of L2 vowels

According to the data analysis in Research article 1 (Jekiel and Malarski 2021), the acquisition of individual L2 vowels varied across participants. The vowels that were produced closer to the model GB pronunciation after training were the front monophthongs TRAP, KIT, and FLEECE (see Wells 1982 for lexical sets). At the same time, the vowels that were the most distanced from the pronunciation model were the back monophthongs FOOT, GOOSE, and LOT, as well as the front DRESS. A possible explanation for these results is the difference between the vowel systems. Compared to twelve monophthongs and eight diphthongs found in GB (Cruttenden 2014), Polish has only six oral and two nasal vowels with no durational differences (Jassem 2003). As a result, Polish learners of English who want to achieve native-like pronunciation have to master a significantly greater number of L2 vowel contrasts than in their L1, including more subtle differences in formant values and vowel duration. In such a case, the acquisition process is often impeded by the assimilation of L2 vowels to L1 categories, as predicted by the Speech Learning Model (Flege 1995).

There is still a limited number of studies on the perception and production of GB vowels by Polish learners of English (e.g. Balas 2018; Porzuczek 2007; Rojczyk 2010; Schwartz and Kaźmierski 2020). First, according to Rojczyk (2011) and Weckwerth (2011), TRAP is usually assimilated by Polish learners of English to /a/ or /ɛ/. This result can be attributed to the “mapping” issue (Bohn 2017) caused by the lack of balance between the two vowel systems, since Polish has only two vowels in the same vowel space where English features three: DRESS, STRUT, and TRAP (Sobkowiak 2008). In Research paper 1 (Jekiel and Malarski 2021), participants produced more native-like GB TRAP after training, which suggests that accent training allowed them to bypass this common problem observed across Polish learners of English.

Next, following Schwartz (2019), a possible explanation why participants successfully acquired KIT and FLEECE during the accent training might stem from the fact that these vowels can be viewed as parallel to Polish /ɨ/ and /i/, since both pairs can be categorised as high front vowels and occupy a similar area on the vowel chart. At the same time, FOOT and GOOSE were the most difficult to acquire, since Polish has only one high back vowel /u/. Additionally, the KIT-FLEECE contrast has a much higher functional load due to high frequency minimal pairs, as opposed to the FOOT-GOOSE contrast (Higgins 2019). According to Bybee (2001), the frequency effect can be a significant factor in the formation of L2 categories, as more common vowel contrasts facilitate L2 vowel acquisition, while less frequent pairs are problematic to acquire. These considerations are also reflected in a perceptual study by Schwartz and Dzierla (2018), in which FOOT and GOOSE were the most difficult to identify by Polish learners of English. These observations are in line with the results found in Research article 1 (Jekiel and Malarski 2021), where KIT and FLEECE were successfully acquired during the accent training, while FOOT and GOOSE remained challenging for the learners.

Finally, the two most difficult vowels to acquire in training were LOT and DRESS. Polish learners of English often assimilate LOT to Polish /ɔ/, which occupies a similar vowel space (Sobkowiak 2008). Similarly for DRESS, it is possible that the Polish learners of English assimilated to Polish /ɛ/ due to the relatively small difference between the two phonemes and the small impact of this process on intelligibility (Rojczyk 2010). Indeed, English DRESS and Polish /ɛ/ are perceptually similar (Balas 2018), especially when accounting for the current lowering of this vowel in GB, as opposed to the raised DRESS in more conservative RP (Hawkins and Midgley 2005; Schmitt 2007). In Research paper 1 (Jekiel and Malarski 2021), we noted that DRESS produced by the pronunciation teachers was much higher than their Polish /ɛ/, making it sound more similar to conservative RP, rather than contemporary GB. Due to language exposure outside the EFL classroom (Szyszka 2018), it is possible that the participants relied on the lowered DRESS found in audio-visual media over the conservative production of this vowel by their pronunciation teachers.

To summarise, the GB vowels considered the most critical for intelligibility in L2 speech and known to be difficult for Polish learners of English (i.e. KIT, FLEECE, and TRAP) were successfully produced as more native-like after the two-semester accent training. At the same time, the GB vowels regarded as less crucial for L2 English comprehensibility and often substituted by Polish learners of English (i.e. FOOT, GOOSE, LOT, DRESS) were still produced as less native-like after training. These conclusions are not only in agreement with former studies on L2 perception and production by Polish EFL learners but are also aligned with the current trend of discussing foreign language pronunciation in terms of intelligibility and comprehensibility, rather than accentedness. The results found in Research article 1 (Jekiel and Malarski 2021) can be viewed as a valuable contribution to the current state of knowledge of L2 speech by Polish learners of English.

2.1.2. The acquisition of L2 rhythm

The results found in Research article 2 (Jekiel 2022) confirmed that the acquisition of L2 rhythm can be observed in a longitudinal study with the use of rhythm metrics. There was apparent progress after training in the production of higher vocalic variation, i.e. participants showed a higher variability in vowel duration and vowel reduction, moving from syllable-timed rhythm towards stress-timed rhythm. The increase in vocalic variation had a direct effect on all post-training rhythm metric scores, which were between the pre-training rhythm metric scores and the rhythm metric scores of the participants' pronunciation teachers. The results of this study are in agreement with previous studies on L2 rhythm, where EFL learners gradually shifted from syllable-timed to stress-timed speech (e.g. Ordin and Polyanskaya 2015; White and Mok 2018). The outcomes of this study are also in line with Gralińska-Brawata (2014), which is the only other study on L2 rhythm of Polish adult learners of English with similar methodology to date. The findings in Research article 2 (Jekiel 2022) confirm that L2 rhythm is teachable and learnable in a formal learning environment during a two-semester accent training.

Language rhythm can be difficult to define and train in the EFL classroom. For that reason, Barry (2007) suggests that teaching L2 rhythm should focus on articulatory practice of the underlying phonological processes affecting perceived speech rhythm, i.e. vowel duration and vowel reduction, since these aspects of pronunciation have a direct effect on English rhythm, especially in GB (White and Mattys 2007), while their insufficient production leads to non-native rhythm (Adams 1979). This teaching method was also applied during the accent training, as pronunciation teachers focused on practising vowels and consonants of English, paying particular attention to vowel duration and vowel reduction.

Consequently, the successful acquisition of these features by Polish learners of English could be observed with the use of rhythm metrics, which calculate those durational differences between consonantal and vocalic intervals.

While some researchers argue that teaching foreign language rhythm is unproductive and should be removed from the EFL classroom, there are a number of studies validating its impact on intelligibility and communicative effectiveness (Levis 2018). Even though L2 rhythm is often affected by L1 prosodic features (Mennen and de Leeuw 2014), the results found in Research article 2 (Jekiel 2022) confirm that the successful acquisition of vowel duration and vowel reduction by Polish advanced learners of English can be considered an improvement in L2 speech rhythm. Hopefully, these findings can be regarded as a worthwhile addition to the research literature in the field of foreign language rhythm.

2.1.3. The acquisition of L2 intonation

Research article 3 (Jekiel and Malarski 2023) revealed an observable progress in the production of L2 intonation after training, i.e. the participants used more native-like intonation patterns, similarly to their pronunciation teachers, after the two-semester accent training course. This finding confirms that L2 intonation can be teachable and learnable in a formal instruction environment. However, the results showed that not all post-training intonation contours were produced correctly. While participants made a noticeable progress in the production of the falling tone in wh-questions and tag questions, the post-training fall-rise tone expressing non-finality and the rise-fall tone expressing strong approval were still the most challenging for the learners. On the one hand, the post-training production of the falling tone in questions shows that the participants managed to move away from the use of a rising tone, which is typical for Polish in these contexts (Mikoś 1976), to a tone more common for GB speakers. On the other hand, the post-training production of the fall-rise and rise-fall tones was still inconsistent. These results suggest that the acquisition of complex L2 intonation patterns, which also do not have an equivalent in L1, can be quite demanding for Polish learners of English, even during a two-semester accent training.

Mastering intonation in the EFL classroom can be a demanding task both in terms of effective learning and teaching (Setter 2008). Foreign language teachers frequently decide to omit this aspect of pronunciation (Demirezen 2009), often due to insufficient teaching materials (Derwing 2008). Many learners of L2 English make errors when producing L2 intonation patterns (Willems 1982), relying on their L1 intonation patterns instead (Gut 2009). In particular, Polish students of English find it difficult to define the notion of intonation and are often unaware of prosody learning strategies (Lewicka-Mroczek and Szymaniuk 2013). At the same time, Polish learners of English are good at discriminating different L2 intonation patterns but struggle at labelling them (Rojczyk and Porzuczek 2017). Nevertheless, intonation is an important aspect of pronunciation, especially in international communication (Aronsson 2014), and should be included in the EFL classroom (as recommended by Chapman 2007). Moreover, recent findings (e.g. Saito 2021) show that successful acquisition of L2 intonation can improve overall L2 intelligibility, suggesting that EFL teachers should prioritise L2 prosody in classroom instruction. To summarise, the results reported in Research paper 3 (Jekiel and Malarski 2023) support the argument that L2 intonation is teachable and learnable in a formal learning environment and can be regarded as a valuable contribution in the vast field of research in foreign language pronunciation.

2.2. Musical hearing and foreign language pronunciation

In all three Research articles (Jekiel and Malarski 2021; Jekiel 2022; Jekiel and Malarski 2023), musical hearing was assessed with the same musical hearing tests designed by Mandell (2009): the Adaptive Pitch Test measuring pitch perception abilities, the Rhythm Test measuring rhythmic memory, and the Tonedeaf Test measuring melodic memory. The results of these tests were then juxtaposed against the pronunciation scores before and after the accent training, i.e. the proximity to the model L2 vowels (Jekiel and Malarski 2021), the proximity to model L2 rhythm (Jekiel 2022), and the proximity to model L2 intonation patterns (Jekiel and Malarski 2023). First, rhythmic memory was reported as a significant predictor of native-like L2 vowel production before the accent training (Jekiel and Malarski 2021). Secondly, only weak positive correlation was found between rhythmic memory and two rhythm metric scores, VarcoV and nPVI-V, with no significant correlations between musical hearing test scores and individual rhythm metric scores before or after training (Jekiel 2022). Finally, pitch perception emerged as a significant predictor of accurate L2 intonation patterns produced after the accent training (Jekiel and Malarski 2023).

The musical hearing tests used in this PhD project were previously tested on over 11,000 subjects (Mandell 2009) and have been recently used in other studies investigating the relationship between musical hearing and foreign language pronunciation. Lorenzen (2019) and Ning (2020) confirmed a correlation between musical hearing, specifically pitch perception and melodic memory, and tonal perception in L2 Mandarin. The tests were also used in studies to investigate the relationship between pitch perception and music reward (e.g. Hernández et al. 2019) or to identify tone deafness (e.g. Mandell et al. 2007; Wang et al. 2023). According to Palomar-García et al. (2020), the musical hearing tests designed by Mandell are reliable and show a good relationship with pitch discrimination, rhythm imitation, and musical abilities across musicians and non-musicians. The following sections discuss in detail the results of each test, as well as the relationship between the musical hearing test scores and specific aspects of foreign language pronunciation.

2.2.1. Pitch perception

The Adaptive Pitch Test was designed to measure pitch perception by playing a series of two tones and asking the participants to determine if the second tone is higher or lower than the first in each pair. Research article 1 (Jekiel and Malarski 2021) found no significant relationship between pitch perception and the acquisition of L2 vowels. At the same time, Research article 3 (Jekiel and Malarski 2023) reported more accurate pitch perception as a significant predictor of more native-like L2 intonation patterns produced by participants after the accent training.

Musical hearing can be defined as inherent sensitivity towards pitch and timing. Both are features of musical sounds, as well as vowels (Kraus et al. 2009), and the link between them was also confirmed in shared processing mechanisms (Lidji et al. 2010). The musical quality of vowels, especially pitch and timbre, is crucial in early language acquisition (Masataka 2007) and plays an important role in prosody of speech and song (Fenk-Oczlon 2017). Despite these commonalities, there is still insufficient research investigating the link between pitch perception and the processing of L2 vowels. For example, Kempe et al. (2015) reported a limited mediating role of pitch perception in the discrimination of non-native vowels and lexical tones. Research article 1 (Jekiel and Malarski 2021) revealed no link between more accurate pitch perception and the production of L2 vowels.

Research article 3 (Jekiel and Malarski 2023) reported pitch perception as a significant predictor of accurate L2 intonation patterns produced after training, i.e. participants who scored better in the Adaptive Pitch Test also produced more native-like L2 intonation patterns after the accent training. Former studies confirmed shared cognitive processes of pitch perception in music and language (e.g. Zatorre and Baum 2012), as well as the effect of musical pitch perception on successful encoding of intonation in speech (e.g. Bidelman and Krishnan 2009), including successful discrimination of non-native intonation patterns (e.g. Dankovicová et al. 2007). A considerable number of studies involving musicians and non-musicians investigated their pitch perception skills in speech (e.g. Magne et al. 2006; Schön et al. 2004), especially in tonal languages (e.g. Alexander et al. 2005; Marie et al. 2010), but also in non-native speech (e.g. Marques et al. 2007; Wong and Perrachione 2007). Moreover, it was found that listeners suffering from tone deafness have problems with discriminating intonation patterns in speech (Patel et al. 2005) and have a phonemic awareness deficit (Loui et al. 2011). The results found in Research article 3 (Jekiel and Malarski 2023) are in agreement with these previous studies and expand the current literature by examining the role of pitch perception in the acquisition of L2 intonation in a formal learning environment.

2.2.2. Rhythmic memory

The Rhythm Test was designed to measure rhythmic memory by playing a series of two short rhythmic patterns and asking the participants to determine if the two patterns were the same or different from one another. Rhythmic memory was reported as a significant predictor of native-like L2 vowel production before the accent training (Jekiel and Malarski 2021), and a moderate predictor of successful acquisition of L2 rhythm (Jekiel 2022).

Rhythm is a fundamental element of language and music (Ding et al. 2017) and it is processed by shared neurocognitive resources (Magne et al. 2016). Consequently, more accurate rhythmic memory was found to play an important role in language acquisition (see Langus et al. 2017 for a review), as well as improved phonological awareness (Moritz et al. 2013). Moreover, since the relationship between speech and music is bidirectional, enhanced musical rhythmic memory can be observed as a result of L2 acquisition (Roncaglia-Denissen et al. 2016). Finally, musical rhythm exercises were found to be an effective tool in the EFL classroom to practise L2 rhythm, especially durational contrasts and syllable stress (Wang et al. 2016), as well as to achieve native-like pronunciation and fluency (Llanes-Coromina et al. 2018). These observations are in agreement with the findings of this PhD project.

In Research article 1 (Jekiel and Malarski 2021), higher scores in the Rhythm Test predicted more native-like production of L2 vowels after the accent training. After closer examination, it was also reported that the most significant result was found for the production of TRAP, both before and after training, suggesting that this aspect of musical hearing can have a strong association with the successful acquisition of this vowel, which is known to be problematic for Polish learners of English. The reason for this may be that while Polish vowels are relatively stable in their quality, English monophthongs are dynamic, i.e. vowel formants shift during their production (Schwartz 2010). Consequently, in order to produce more native-like L2 vowels, Polish learners of English need to temporally reorganise the production of their vocalic targets (Schwartz and Kaźmierski 2020). Since this process can be viewed as closely related to rhythmic memory, it might be a plausible explanation why participants with higher Rhythm Test scores produced more native-like L2 vowels. More studies investigating the role of rhythmic memory in the acquisition of vowel inherent spectral change are required in order to reinforce this claim.

In Research article 2 (Jekiel 2022), a weak positive correlation was reported between the Rhythm Test scores and the rhythm metric scores after training, i.e. participants with more accurate rhythmic memory also produced more native-like L2 vowel duration and vowel reduction. Since these aspects can be regarded as segmental features and language rhythm features, their accurate production in L2 speech can be perceived as successful acquisition of L2 rhythm. The findings of this study can be a noteworthy supplement to Cason et al. (2020), in which musical rhythmic skills predicted more accurate production of L2 prosody.

2.2.3. Melodic memory

The Tonedeaf Test was designed to measure melodic memory by presenting sets of two short instrumental melodies and asking the participants to determine if the two melodies were the same or different from one another. The test was used in all three Research articles

(Jekiel and Malarski 2021; Jekiel 2022; Jekiel and Malarski 2023) to investigate the relationship between melodic memory and the production of L2 vowels, L2 rhythm, and L2 intonation patterns. No significant results were found for this particular musical hearing test. These results differ from the recent studies using the same musical hearing test (Lorenzen 2019; Ning 2020), which reported a significant relationship between melodic memory and pitch discrimination in L2 tone languages.

A melody can be generally defined as a sequence of musical tones, or a succession of pitches in rhythm, and makes up the most memorable aspect of music. For this reason, melodic memory is a multifaceted skill related to pitch perception, rhythm perception, and working memory. Recent research suggests that music processing and speech processing rely on similar working memory resources, e.g. Fennell et al. (2021) reported that musicians who were better at discriminating between two musical melodies were also better at language processing. This PhD project expected to observe a similar link between the results from the Tonedeaf Test assessing melodic memory and the production of L2 speech. However, while rhythmic memory emerged as a significant predictor of native-like L2 vowel production, and pitch perception was a significant predictor of native-like L2 intonation, no significant result was found for melodic memory. It is possible that the complex nature of melodic memory and its reliance on different aspects of musical hearing skills could not be evidently linked with specific elements of L2 pronunciation. Moreover, pitch range and pitch variability are used significantly more in musical melodies than in the speech signal, where pitch range is relatively limited (Chow and Brown 2018). More studies are required to investigate the connection between melodic memory and the acquisition of L2 speech, especially among non-tonal languages.

2.3. Musical experience and foreign language pronunciation

Apart from investigating the influence of musical hearing on foreign-language pronunciation, musical experience was also included as a separate factor affecting the acquisition of L2 speech in all three Research articles (Jekiel and Malarski 2021; Jekiel 2022; Jekiel and Malarski 2023). Pre-existing musical experience was assessed with a questionnaire, asking participants whether they had attended a music school or acquired any particular musical experience before the accent training, including singing or playing a musical instrument, and if so, how long their experience was. The self-reported musical experience was then juxtaposed against the L2 pronunciation data. Research article 1 (Jekiel and Malarski 2021) reported musical experience as a significant factor in the post-training production of L2 vowels, while Research article 2 (Jekiel 2022) and Research article 3 (Jekiel and Malarski 2023) found no significant relationship between musical experience and the production of L2 rhythm or L2 intonation.

Previous studies investigating the link between musical training and L2 pronunciation suggest the potential influence of musical practice on the acquisition of L2 vowels (e.g. Intartaglia et al. 2017; Götz et al. 2023) and L2 prosody (e.g. Dankovicová et al. 2007; Marques et al. 2007), as well as phonological awareness (e.g. Moritz et al. 2013) and speech imitation (Pastuszek-Lipińska 2008). By analogy, the current PhD thesis expected to observe similar parallels between musical experience and the acquisition of L2 pronunciation in a formal learning environment. The results suggest that former musical practice can predict more native-like production of L2 vowels after a two-semester accent training. This relationship can be explained due to the established acoustic similarities between sounds in music and vowels in speech, as well as the shared processing mechanisms (Kempe et al. 2015). As a result, musicians should have more accurate L2 speech perception, which in turn should allow them to produce more accurate L2 vowels after the accent training.

This PhD project was primarily concerned with the relationship between foreign-language pronunciation and pre-existing musical hearing skills, rather than formal music education or informal musical practice. There were two main reasons for establishing this focal point. First, there is an observable dominance of studies investigating the role of musical training in L2 acquisition, while musical hearing skills of participants without formal or amateur musical experience are often omitted. Current studies suggest that pre-existing musical hearing may be more conducive to the acquisition of non-native pronunciation than former musical training (see Lund 2023 for a review). For example, in a recent study by Nisha et al. (2021), non-musicians with good musical hearing skills can be similar to musicians in terms of enhanced selective attention skills and working memory. Therefore, both aspects should be taken into account when investigating the factors affecting L2 speech. Second, the participants of this PhD project were undergraduate students of English and their prior musical experience varied in terms of its specificity and duration (e.g. only four participants had attended music school prior to their studies). Overall, only a third of participants reported some musical experience, including singing or playing a musical instrument in various contexts. For this reason, the number of self-reported years of musical experience was the only variable included in the analysis. There is a general agreement in literature that six years of musical practice is the threshold from which participants can be classified as musicians (see Zhang et al. 2020 for a review). This cut-off point was also an indicator during the questionnaire data screening for this PhD project, as self-reported musical experience of shorter periods of time was not considered in the data analysis.

2.4. Research limitations and future directions

This PhD project produced new evidence for the influence of musical hearing in the acquisition of foreign language pronunciation by conducting a multi-modal longitudinal study investigating the role of pitch perception, rhythmic memory, and melodic memory in the production of non-native vowels, rhythm, and intonation during a two-semester accent training. Despite such a comprehensive approach, the discussion of the reported findings revealed a number of limitations, as well as directions for future research.

First of all, a single source of participants, who were all Polish undergraduate students of English of a state university, precluded the possibility to have a control group with no accent training or to control for the participants' prior musical experience. The practical English pronunciation course and the theoretical English phonetics and phonology course were obligatory for all students during the first two semesters of English studies. Consequently, it was impossible to have a control group comprising a similar pool of participants without the accent training. Furthermore, there was no possibility to have a balanced sample of participants in terms of prior music education or other types of former musical experience. As a result, the primary objective of this PhD project was to measure the influence of musical hearing in the acquisition of L2 pronunciation during the accent training, while musical experience was kept as a secondary objective. Future studies should consider not only a balanced sample in terms of the participants' musical background, but also include students of English from various universities to investigate the effect of accent training alone.

Secondly, the phonetic data collection relied on read-aloud tasks, rather than spontaneous speech. The controlled tasks utilised a monitor to present a series of monosyllabic words, an elicitation paragraph, and a set of short dialogues. These were read by participants to elicit L2 vowels, rhythm, and intonation consecutively. While this procedure may be considered as standard practice in phonetic research, there is still a scarcity of studies investigating natural L2 speech. According to Chau et al. (2022), exploring improvements in spontaneous speech, as opposed to controlled speech, should be considered as more important, since the ability to speak casually and fluently in L2 is often the primary goal for foreign language learners. Moreover, previous studies point to the influence of orthographic prompts in the production of non-native sounds (e.g. Kato and Baese-Berk 2020). This is directly related to the spelling pronunciation problem, which leads to incorrect pronunciation, and is often found to be a major issue for Polish learners of English (e.g. Nowacka 2016). Therefore, the inclusion of spontaneous speech in this type of research would not only allow to observe the difference between controlled and free speech, but also to investigate the relationship between musical hearing and different types of speech data.

Thirdly, while the use of the musical hearing tests designed by Mandell (2009) was a relatively novel approach, applying other measures of musical aptitude could provide results that are more comparable to earlier studies. However, many previous musical aptitude tests (e.g. Gordon 1989; Seashore et al. 1960) have been criticised for their length, difficulty, and inaccessibility (see Law and Zentner 2012 for a review). On the other hand, there are also new developments taking place in terms of musical aptitude assessment, e.g. the current iteration of the online Profile of Music Perception Skills, Micro-PROMS (Strauss et al. 2023) is easily accessible and allows to objectively test musical hearing skills in circa 10 minutes (for comparison, the original Full-PROMS requires circa 60 minutes). The implementation of such a tool would allow to focus on gathering linguistic data in phonetic research and provide sufficient evaluation of participants' musical aptitude at the same time. To summarise, there is a growing need to objectively assess musical skills in an accessible way, so that future studies investigating the influence of musical hearing on foreign language pronunciation can share similar methodology and have comparable test results.

Finally, in spite of the multi-modal approach in the present PhD project, there are other factors predicting learners' success in the acquisition of L2 pronunciation that should be taken into account in future studies, including language talent (e.g. Dogil and Reiterer 2009), anxiety (e.g. Baran-Łucarz 2013) and motivation (e.g. Smit 2002). In particular, advanced students of English may be highly motivated to acquire native-like pronunciation (Nowacka 2012). Although all participants received similar in-class pronunciation practice, it is possible that more motivated students spent more time on self-study and consequently produced more native-like targets after the accent training, regardless of their pre-existing musical hearing skills or prior musical experience. In other words, motivation can be a factor allowing to compensate for any deficiencies in terms of musical aptitude or musical training.

Overall, the concept of this PhD project was based on the interplay between perception and production. While the primary objective of all three Research articles (Jekiel and Malarski 2021; Jekiel 2022; Jekiel and Malarski 2023) was to observe the influence of musical hearing on foreign language pronunciation, it is also possible that the accent training had an influence on participants' musical hearing skills. Previous research reported that production can affect perception during foreign language learning (e.g. Baese-Berk 2019) and learning a second language can even have an impact on some musical hearing skills (e.g. Roncaglia-Denissen et al. 2016). Indeed, the practical English pronunciation course involved a series of activities targeted not only at production, but also perception, which might have affected the learners' musical skills after the accent training. Since the musical hearing tests were only administered before the accent training, future studies might aim to investigate further the complex interactions between musical hearing and the acquisition of foreign language pronunciation.

Conclusion

The present PhD project explored the influence of musical hearing and musical experience on the acquisition of foreign language pronunciation in Polish advanced learners of English. The longitudinal phonetic study, which was reported in three Research articles (Jekiel and Malarski 2021; Jekiel 2022; Jekiel and Malarski 2023) and summarised in this PhD thesis, provides new evidence for the role of pitch perception, rhythmic memory, and melodic memory, as well as former musical experience, in the successful acquisition of L2 vowels, rhythm, and intonation in a formal learning environment.

First of all, observable post-training improvement in non-native pronunciation was reported across all three Research articles. In Research article 1 (Jekiel and Malarski 2021), participants produced more native-like L2 vowels after training, as the mean Euclidean distance between the post-training vowel formants and the pronunciation teachers' model vowels significantly decreased when compared to the pre-training results. The progress was more apparent for high front vowels than high back vowels, demonstrating that some L2 vowel contrasts may be easier to acquire in a formal learning environment, especially if they have a higher functional load and are vital for intelligibility (Schwartz 2019). In Research article 2 (Jekiel 2022), participants' L2 rhythm metric scores changed significantly as a result of the accent training, which indicates that vowel duration contrasts and vowel reduction were more frequently applied in post-training L2 speech. This finding is in agreement with previous research reporting a gradual shift from syllable-timed to stress-timed rhythmic patterns across L2 learners with different L1 rhythm (Gralińska-Brawata 2014). In Research article 3 (Jekiel and Malarski 2023), participants produced more native-like L2 intonation patterns after training, similarly to their pronunciation teachers. This finding is in line with previous arguments that intonation is both teachable and learnable

(Aronsson 2014). However, the progress was less apparent for fall-rise and rise-fall patterns, which may be considered as difficult for Polish learners of English.

Secondly, the influence of specific aspects of musical hearing on selected elements of non-native pronunciation was also reported across all three studies. In Research article 1 (Jekiel and Malarski 2021), participants with higher rhythmic memory test scores produced more native-like L2 vowels before the accent training. This finding demonstrates that accurate perception of musical rhythm can be a predictor of accurate L2 vowel production and the temporal reorganisation of L2 vocalic targets, which is essential for Polish EFL learners to achieve native-like English pronunciation (Schwartz and Kaźmierski 2020). In Research article 2 (Jekiel 2022), a weak positive correlation was found between the rhythm metric scores and the rhythmic memory test results. This result supplements previous studies investigating the role of musical rhythm perception in L2 speech (Cason et al. 2020). Future studies should include other aspects of language rhythm when investigating the relationship between musical rhythm and non-native prosody. In research article 3 (Jekiel and Malarski 2023), pitch perception was a significant predictor in the acquisition of L2 intonation after training. This finding is in agreement with previous studies investigating musical aptitude and foreign language intonation (e.g. Dankovicová et al. 2007; Marques et al. 2007).

Finally, the current PhD project found only limited influence of musical experience on the acquisition of L2 speech. In Research article 1 (Jekiel and Malarski 2021), former musical experience was a significant predictor on the acquisition of L2 vowels during the accent training. However, no link between musical practice and L2 rhythm (Jekiel 2022) or L2 intonation (Jekiel and Malarski 2023) was observed. This finding is in line with recent studies suggesting that pre-existing musical hearing skills are more conducive in foreign language acquisition than former musical training (see Lund 2023 for a review). Further research incorporating a balanced sample of musicians and non-musicians with similar L2 proficiency is needed to corroborate these results.

Overall, the findings summarised in this PhD thesis and reported across the three Research articles (Jekiel and Malarski 2021; Jekiel 2022; Jekiel and Malarski 2023) provide new evidence and offer a novel insight into research on musical hearing and musical experience in foreign language pronunciation, suggesting that pitch perception and rhythmic memory can have an influence on the acquisition of non-native segments and prosody in a formal learning environment. Nevertheless, more research is still required in order to formulate a better understanding of the complex interplay between musical skills and foreign language acquisition, including new methods of assessing musical hearing and non-native speech, as well as other factors affecting the acquisition of foreign language pronunciation.

Abstract

Research in language and music is a broad field, ranging from interdisciplinary works on language and music evolution (Brown 2000) to neurolinguistics (Patel 2008). Over the years, studies exploring the interplay between the two domains have identified a number of shared attributes, including acoustic features and processing mechanisms (Slevc 2012). These relationships between music and speech led researchers to examine the influence of musical aptitude on language acquisition (Brandt et al. 2012), including foreign language speech (Chobert and Besson 2013). While the link between musical hearing and speech perception has been widely investigated (e.g. Marques et al. 2007; Nardo and Reiterer 2009), the relationship between musical skills and non-native speech production has still not received enough attention (e.g. Dolman and Spring 2014; Milovanov et al. 2010). Moreover, the sheer variability of methods leaves room for further research, as many former studies relied on self-reported foreign language proficiency, impressionistic assessment of speech, or self-reported musical training. Finally, there is still no longitudinal study that would measure the influence of specific musical skills on individual elements of foreign language pronunciation in a formal learning environment.

To address this research gap, the present PhD project aims to investigate the influence of musical hearing and musical experience on the acquisition of foreign language pronunciation in Polish advanced learners of English. To this end, a longitudinal phonetic study was conducted among 50 Polish undergraduate students of English attending a two-semester accent training, utilising acoustic speech analysis and musical hearing assessment. The results were reported across three Research articles investigating the influence of pre-existing musical hearing skills and former musical experience on L2 vowel quality (Jekiel

and Malarski 2021), L2 vowel duration and reduction affecting language rhythm (Jekiel 2022), and L2 intonation patterns (Jekiel and Malarski 2023).

Research article 1 (Jekiel and Malarski 2021) reported a study investigating whether pre-existing musical hearing skills and former musical experience can predict successful acquisition of L2 vowels in a formal learning environment. Participants were recorded when reading aloud a series of monosyllabic words to elicit GB vowels before and after a two-semester accent training. To assess their musical hearing and musical experience, participants completed three musical hearing tests measuring their pitch perception, rhythmic memory, and melodic memory, followed by a questionnaire on their former musical experience. The results revealed that rhythmic memory can be a significant predictor of native-like production of L2 vowels before training, while years of musical experience can be a significant predictor of accurate L2 vowel production after training. While pre-existing rhythmic memory and former musical experience can have a positive influence on the successful acquisition of foreign language vowels, former pronunciation skills remained as the deciding factor in the acquisition of L2 speech in a formal learning environment.

Research article 2 (Jekiel 2022) investigated the influence of musical hearing and musical experience on native-like L2 rhythm production before and after a two-semester accent training. The findings reported a significant difference in the rhythm metric scores for vocalic intervals, demonstrating participants' higher variation in vowel duration and vowel reduction in controlled speech after training. The results confirmed that rhythm metrics can be successfully used to observe progress in L2 rhythm production in a formal learning environment. A weak correlation emerged between the rhythm metric scores and rhythmic memory test scores, suggesting a potential link between musical hearing and the acquisition of L2 rhythm among advanced learners of English.

Research article 3 (Jekiel and Malarski 2023) examined the influence of musical hearing and musical experience on native-like production of L2 intonation patterns after a two-semester accent training. The results revealed that accurate pitch perception can be a significant predictor of native-like L2 intonation after training. At the same time, musical experience was not found to be a significant factor in the acquisition of foreign language intonation. While higher pitch perception test scores were correlated with more correct production of L2 intonation patterns after the accent training, participants without superior musical hearing or previous musical experience also improved, indicating that L2 pronunciation training in a formal learning environment can lead to a significant progress in native-like production of L2 intonation.

Overall, the studies reported in the three Research articles (Jekiel and Malarski 2021; Jekiel 2022; Jekiel and Malarski 2023) and summarised in this PhD thesis provide new evidence for the role of musical hearing and musical experience in foreign language pronunciation; specifically, that pitch perception and rhythmic memory can have an influence on the acquisition of non-native segments and prosody in a formal learning environment.

Streszczenie

Badania nad językiem i muzyką stanowią szeroką dziedzinę, obejmującą interdyscyplinarne prace z zakresu zarówno ewolucji języka i muzyki (Brown 2000), jak i neurolingwistyki (Patel 2008). Przez lata badania zgłębiające wzajemne oddziaływanie tych dwóch sfer zidentyfikowały szereg wspólnych atrybutów, w tym cechy akustyczne i mechanizmy przetwarzania (Slevc 2012). Te związki między muzyką i mową skłaniają naukowców do badania wpływu uzdolnień muzycznych na przyswajanie języka (Brandt i in. 2012), w tym mowy języka obcego (Chobert i Besson 2013). Chociaż związek między słuchem muzycznym a percepcją mowy został szeroko zbadany (np. Marques i in. 2007; Nardo i Reiterer 2009), zależność między umiejętnościami muzycznymi a produkcją mowy w języku obcym nadal nie otrzymała wystarczającej uwagi (np. Dolman i Spring 2014; Milovanov i in. 2010). Co więcej, sama różnorodność metod pozostawia miejsce na kolejne badania, jako że wiele badań opierało się na samodzielnie zadeklarowanej znajomości języka obcego i edukacji muzycznej czy impresjonistycznej ocenie wymowy. Wciąż nie powstało zbyt wiele badań podłużnych, które mogłyby zmierzyć wpływ określonych umiejętności muzycznych na poszczególne elementy wymowy języka obcego w formalnym środowisku edukacyjnym.

Aby wypełnić tę lukę badawczą, niniejszy projekt doktorski zgłębia wpływ słuchu muzycznego i doświadczenia muzycznego na przyswajanie wymowy języka obcego przez polskich uczniów języka angielskiego na poziomie zaawansowanym. W tym celu przeprowadzono podłużne badanie fonetyczne na 50 polskich studentach filologii angielskiej uczęszczających na dwusemestralny kurs akcentu brytyjskiego. Badanie wykorzystuje akustyczną analizę mowy i ocenę słuchu muzycznego, a jego wyniki zawarto w trzech artykułach omawiających wpływ słuchu muzycznego i doświadczenia muzycznego na

jakość samogłosek w języku drugim (Jekiel i Malarski 2021), czas trwania i redukcję samogłosek jako czynników kształtujących rytm języka (Jekiel 2022) oraz kontury intonacyjne (Jekiel i Malarski 2023).

Artykuł 1 (Jekiel i Malarski 2021) opisuje badanie mające na celu sprawdzenie czy umiejętności muzyczne i wcześniejsze doświadczenie muzyczne mogą pomyślnie wpływać na przyswajanie samogłosek języka drugiego w formalnym środowisku edukacyjnym. Uczestnicy zostali zarejestrowani podczas czytania na głos serii jednosylabowych słów w celu pozyskania nagrań samogłosek brytyjskich przed i po dwusemestralnym treningu akcentu. Aby ocenić słuch muzyczny i doświadczenie muzyczne, uczestnicy wykonali trzy testy ze słuchu muzycznego mierzące ich percepcję tonalną, pamięć rytmiczną oraz pamięć melodyczną, a następnie wypełnili kwestionariusz dotyczący ich wcześniejszych doświadczeń muzycznych. Wyniki pokazały, że pamięć rytmiczna może być istotnym predykatorem niemal natywnej w jakości produkcji samogłosek języka drugiego przed treningiem, a lata doświadczenia muzycznego mogą przewidzieć dokładną produkcję samogłosek języka drugiego po treningu. Podczas gdy pamięć rytmiczna i wcześniejsze doświadczenia muzyczne mogą pozytywnie wpływać na pomyślne opanowanie samogłosek w języku obcym, początkowe umiejętności z zakresu wymowy były decydującym czynnikiem w przyswajaniu mowy języka drugiego w formalnym środowisku edukacyjnym.

Artykuł 2 (Jekiel 2022) badał wpływ słuchu muzycznego i doświadczeń muzycznych na produkcję rytmu języka drugiego przed i po dwusemestralnym treningu akcentu. Wyniki pokazały znaczną różnicę w wynikach metryk rytmu dla interwałów wokalnych, wykazując większe zróżnicowanie czasu trwania samogłosek i redukcji samogłosek w kontrolowanej mowie po treningu. Wyniki potwierdziły, że metryki rytmu mogą z powodzeniem być wykorzystywane do obserwacji postępów w przyswajaniu rytmu języka drugiego w formalnym środowisku edukacyjnym. Słaba pozytywna korelacja między wynikami metryk rytmu a wynikami testu pamięci rytmicznej może sugerować potencjalny związek między słuchem muzycznym a przyswajaniem rytmu języka drugiego przez uczniów języka angielskiego na poziomie zaawansowanym.

Artykuł 3 (Jekiel i Malarski 2023) badał wpływ słuchu muzycznego i doświadczeń muzycznych na produkcję konturów intonacyjnych języka drugiego po dwusemestralnym treningu akcentu. Wyniki pokazały, że precyzyjne postrzeganie wysokości tonu może być istotnym predykatorem niemal natywnej intonacji języka drugiego po treningu. Jednocześnie nie stwierdzono, aby doświadczenie muzyczne było istotnym czynnikiem w przyswajaniu intonacji języka obcego. Mimo że wyniki testu percepcji tonów były skorelowane z bardziej poprawnym wytwarzaniem konturów intonacji języka drugiego po treningu, uczestnicy bez lepszego słuchu muzycznego lub wcześniejszego doświadczenia muzycznego również zrobili postępy, co wskazuje, że trening wymowy w formalnym środowisku edukacyjnym może prowadzić do znacznego postępu w produkcji intonacji języka drugiego.

Podsumowując, badania przedstawione w trzech artykułach badawczych (Jekiel i Malarski 2021; Jekiel 2022; Jekiel i Malarski 2023) i omówione w niniejszej rozprawie doktorskiej dostarczają nowych dowodów na znaczenie słuchu i doświadczenia muzycznego w przyswajaniu wymowy języka obcego; w szczególności na to, że percepcja tonalna i pamięć rytmiczna mogą wpływać na przyswajanie segmentów oraz prozodii języka drugiego w kontekście formalnego nauczania.

References

Adams, Corinne. 1979. *English speech rhythm and the foreign learner*. The Hague: Moulton.

Alexander, Jennifer A., Patrick C. M. Wong and Ann R. Bradlow. 2005. "Lexical tone perception in musicians and non-musicians", in: Gunilla Svanfeldt and Dirk Olszewski (eds.), *Proceedings of Interspeech 2005, Eurospeech, 9th European Conference on Speech Communication and Technology.* Lisbon: International Speech Communication Association, 397-400.

Anvari, Sima H., Laurel J. Trainor, Jennifer Woodside and Betty A. Levy. 2002. "Relations among musical skills, phonological processing, and early reading ability in preschool children", *Journal of Experimental Child Psychology* 83, 2: 111-130.

Aronsson, Berit. 2014. "Prosody in the foreign language classroom - always present, rarely practised?", *Journal of Linguistics and Language and Teaching* 5: 207-225.

Arvaniti, Amalia. 2012. "The usefulness of metrics in the quantification of speech rhythm", *Journal of Phonetics* 40, 3: 351-373.

Baese-Berk, Melissa Michaud. 2019. "Interactions between speech perception and production during learning of novel phonemic categories", *Attention, Perception, and Psychophysics* 81, 4: 981-1005.

Balas, Anna. 2018. "English vowel perception by Polish advanced learners of English", *Canadian Journal of Linguistics* 63, 3: 309-338.

Baran-Łucarz, Małgorzata. 2013. "Phonetics learning anxiety: Results of a preliminary study", *Research in Language* 11: 57-79.

Barry, William J. 2007. “Rhythm as an L2 problem. How prosodic is it?”, in: Jürgen Trouvain and Ulrike Gut (eds.), *Non-native prosody: Phonetic description and teaching practice.* New York: De Gruyter Mouton, 97-120.

Beckman, Mary E. and Gayle A. Elam. 1997. *Guidelines for ToBI labeling.* Columbus: The Ohio State University.

Bella, Simone D., Nicolas Farrugia, Charles-Etienne Benoit, Valentin Begel, Laura Verga, Eleanor Harding and Sonja A. Kotz. 2017. “BAASTA: Battery for the assessment of auditory sensorimotor and timing abilities”, *Behavior Research Methods* 49: 1128-1145.

Bentley, Arnold. 1969. “Measurement and development of musical abilities”, *Journal of Research in Music Education* 17: 41-46.

Besson, Mireille, Julie Chobert and Céline Marie. 2011. “Transfer of training between music and speech: Common processing, attention, and memory”, *Frontiers in Psychology* 2: 94.

Besson, Mireille, Daniele Schön, Sylvain Moreno, Andréia Santos and Cyrille Magne. 2007. “Influence of musical expertise and musical training on pitch processing in music and language”, *Restorative Neurology and Neuroscience* 25, 3-4: 399-410.

Bidelman, Gavin M. and Ananthanarayan Krishnan. 2009. “Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem”, *Journal of Neuroscience* 29, 42: 13165-13171.

Boersma, Paul and David Weenink. 2023. Praat: doing phonetics by computer [Computer program]. Version 6.3.10. (http://www.praat.org/) (date of access: 1 Jun. 2023).

Bohn, Ocke-Schwen. 2017. “Cross-language and second language speech perception.”, in: Eva M. Fernandez and Helen S. Cairns (eds.), *The handbook of psycholinguistics*. New York: Wiley, 213-239.

Borodkin, Katy, Tamar Gassner, Hadeel Ershaid and Noam Amir. 2022. “tDCS modulates speech perception and production in second language learners”, *Scientific Reports* 12: 16212.

Brandt, Anthony, Molly Gebrian and L. Robert Slevc. 2012. “Music and early language acquisition”, *Frontiers in Psychology* 3: 327.

Brown, Steven. 2000. “The ‘musilanguage’ model of language evolution”, in: Nils L. Wallin, Merker Björn and Steven Brown (eds.), *The origins of music*. Cambridge: MIT Press, 271-300.

Brown, Steven, Michael J. Martinez and Lawrence. M. Parsons. 2006. “Music and language side by side in the brain: A PET study of the generation of melodies and sentences”, *The European Journal of Neuroscience* 23: 2791-2803.

Burri, Michael. 2023. “Comparing L2 teachers’ practices with learners’ perceptions of English pronunciation teaching”, *Profile: Issues in Teachers’ Professional Development* 25, 1: 129-145.

Bybee, J. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.

Cason, Nia, Muriel Marmursztejn, Mariapaola D’Imperio and Daniele Schön. 2020. “Rhythmic abilities correlate with L2 prosody imitation abilities in typologically different languages”, *Language and Speech* 63, 1: 149-165.

Chapman, Mark. 2007. “Theory and practice of teaching discourse intonation”, *ELT Journal* 61, 1: 3-11.

Chau, Tue, Amanda Huensch, Yen K. Hoang and Hiep T. Chau. 2022. “The effects of L2 pronunciation instruction on EFL learners’ intelligibility and fluency in spontaneous speech”, *Teaching English as a Second Language Electronic Journal* 25, 4.

Chobert, Julie and Mireille Besson. 2013. “Musical expertise and second language learning”, *Brain Sciences* 3, 2: 923-940.

Choi, William. 2022. “What is ‘music’ in music-to-language transfer? Musical ability but not musicianship supports Cantonese listeners’ English stress perception”, *Journal of Speech, Language, and Hearing Research* 65, 11: 4047-4059.

Chow, Ivan and Steven Brown. 2018. “A musical approach to speech melody”, *Frontiers in Psychology* 9: 1-17.

Christiner, Markus and Susanne M. Reiterer. 2016. “Music, song and speech: A closer look at the interfaces between musicality, singing and individual differences in phonetic language aptitude”, in: Gisela Granena, Daniel O. Jackson and Yucel Yilmaz (eds.), *Cognitive individual differences in second language processing and acquisition.* Amsterdam: John Benjamins, 131-156.

Cruttenden, Alan. 2014. *Gimson’s pronunciation of English*. (8th edition.) London: Routledge.

Culp, Mara E. 2017. "The relationship between phonological awareness and music aptitude", *Journal of Research in Music Education* 65, 3: 328-346.

Cumming, Ruth E. 2011. "The language-specific interdependence of tonal and durational cues in perceived rhythmicality", *Phonetica* 68: 1-25.

Dankovicová, Jana, Jill House, Anna Crooks and Katie Jones. 2007. "The relationship between musical skills, music training, and intonation analysis skills", *Language and Speech* 50, 2: 177-225.

Darwin, Charles. 1871. *The descent of man, and selection in relation to sex*. London: John Murray.

Dege, Franziska and Gudrun Schwarzer. 2011. "The effect of a music program on phonological awareness in preschoolers", *Frontiers in Psychology* 2, 124.

Deguchi, Chizuru, Magali Boureux, Michela Sarlo, Mireille Besson, Massimo Grassi, Daniele Schön and Lucia Colombo. 2012. "Sentence pitch change detection in the native and unfamiliar language in musicians and non-musicians: Behavioral, electrophysiological and psychoacoustic study", *Brain Research* 1455: 75-89.

Delogu, Franco and Yi Zheng. 2020. "Beneficial effects of musicality on the development of productive phonology skills in second language acquisition", *Frontiers in Neuroscience* 14, 618.

Demirezen, Mehmet. 2009. "An analysis of the problem-causing elements of intonation for Turkish teachers of English", *Procedia – Social and Behavioral Sciences* 1, 1: 2776-2781.

Derwing, Tracey M. 2008. "Curriculum issues in teaching pronunciation to second language learners" in: Jette G. Hansen Edwards and Mary L. Zampini (eds.), *Phonology and second language acquisition*. Amsterdam: John Benjamins, 347-369.

Derwing, Tracey M. and Murray J. Munro. 2013. "The development of L2 oral language skills in two L1 groups: A 7-year study", *Language Learning* 63, 2: 163-185.

Derwing, Tracey M., Murray J. Munro, Jennifer A. Foote, Erin Waugh and Jason Fleming. 2014. "Opening the window on comprehensible pronunciation after 19 years: A workplace training study", *Language Learning* 64, 3: 526-548.

Dewaele, Jean-Marc and James McCloskey. 2015. "Attitudes towards foreign accents among adult multilingual language users", *Journal of Multilingual and Multicultural Development* 36, 3: 221-238.

Ding, Nai, Aniruddh D. Patel, Lin Chen, Henry Butler, Cheng Luo and David Poeppel. 2017. "Temporal modulations in speech and music", *Neuroscience and Biobehavioral Reviews* 81: 181-187.

Dogil, Grzegorz and Susanne M. Reiterer (eds.). 2009. *Language talent and brain activity*. New York: De Gruyter Mouton.

Dolman, Matthew and Ryan Spring. 2014. "To what extent does musical aptitude influence foreign language pronunciation skills? A multi-factorial analysis of Japanese learners of English", *World Journal of English Language* 4, 4: 1-11.

Doughty, Catherine J. 2019. "Cognitive language aptitude", *Language Learning* 69: 101-126.

Draxler, Christoph and Klaus Jänsch. 2004. "SpeechRecorder – a Universal Platform Independent Multi-Channel Audio Recording Software", in: Maria Teresa Lino, Maria Francisca Xavier, Fátima Ferreira, Rute Costa and Raquel Silva (eds.), *Proceedings of the Fourth International Conference on Language Resources and Evaluation*. Lisbon: European Language Resources Association, 559-562.

Dufour, Sophie and Noël Nguyen. 2013. "How much imitation is there in a shadowing task?", *Frontiers in Psychology* 4: 346.

Fadiga, Luciano, Laila Craighero and Alessandro D'Ausilio. 2009. "Broca's area in language, action, and music", *Annals of the New York Academy of Sciences* 1169: 448-458.

Fenk-Oczlon, Gertraud. 2017. "What vowels can tell us about the evolution of music", *Frontiers in Psychology* 8: 1581.

Fennell, Anna Marie, Jennifer A. Bugos, Brennan R. Payne and Elizabeth R. Schotter. 2021. "Music is similar to language in terms of working memory interference", *Psychonomic Bulletin and Review* 28: 512-525.

Fitch, W. Tecumseh. 2010. *The evolution of language*. New York: Cambridge University Press.

Flaugnacco, Elena, Luisa Lopez, Chiara Terribili, Marcella Montico, Stefania Zoia and Daniele Schön. 2015. "Music training increases phonological awareness and reading skills in developmental dyslexia: A randomized control trial", *PLoS ONE* 10, 9: e0138715.

Flege, James Emil. 1995. "Second-language speech learning: Theory, findings, and problems", in: Winifred Strange (ed.), *Speech perception and linguistic experience: Issues in cross-language research*. Timonium: York Press, 229-273.

Flege, James Emil and Ocke-Schwen Bohn. 2021. "The revised speech learning model (SLM-r)", in: Ratree Wayland (ed.), *Second language speech learning: Theoretical and empirical progress*. Cambridge: Cambridge University Press, 3-83.

Fuchs, Robert. 2014. "Integrating variability in loudness and duration in a multidimensional model of speech rhythm: Evidence from Indian English and British English", in: Nick Campbell, Dafydd Gibbon and Daniel Hirst (eds.), *Proceedings of the 7th International Conference on Speech Prosody 2014.* Dublin: International Speech Communication Association, 290-294.

Galves, Antonio, Jesus Garcia, Denise Duarte and Charlotte Galves. 2002. "Sonority as a basis for rhythmic class discrimination", in: Bernard Bel and Isabelle Marlien (eds.), *Proceedings of Speech Prosody 2002.* Aix-en-Provençe: Laboratoire Parole et Langage, 323-326.

Gordon, Edwin E. 1989. *Advanced measures of music audiation*. Chicago: GIA.

Gottfried, Terry L. 2007. "Music and language learning", in: Ocke-Schwen Bohn and James J. Munro (eds.), *Language experience in second language speech learning: In honor of James Emil Flege.* Amsterdam: John Benjamins, 221-238.

Götz, Antonia, Liquan Liu, Barbara Nash and Denis Burnham. 2023. "Does musicality assist foreign language learning? Perception and production of Thai vowels, consonants and lexical tones by musicians and non-musicians", *Brain Sciences* 13, 5: 810.

Grabe, Esther and Ee Ling Low. 2002. "Durational variability in speech and the rhythm class hypothesis", *Laboratory Phonology* 7: 515-546.

Gralińska-Brawata, Anna. 2014. "Language experience and phonetic training as factors influencing timing organisation in Polish learners of English", *Research in Language* 12, 2: 185-198.

Gralińska-Brawata, Anna and Paulina Rybińska. 2017. "The relationship between the production of word stress and musical abilities in Polish learners of English", *Research in Language* 15, 3: 265-283.

Gut, Ulrike. 2009. *Non-native speech: A corpus-based analysis of phonological and phonetic properties of L2 English and German*. Frankfurt: Peter Lang.

Gut, Ulrike. 2012. “Rhythm in L2 speech”, in: Dafydd Gibbon, Daniel Hirst and Nick Campbell (eds.), *Rhythm, melody and harmony in speech: Studies in honour of Wiktor Jassem, special issue of speech and language technology 14/15.* Poznań: Polskie Towarzystwo Fonetyczne, 83-94.

Hawkins, Sarah and Jonathan Midgley. 2005. “Formant frequencies of RP monophthongs in four age groups of speakers”, *Journal of the International Phonetic Association* 35, 2: 183-199.

Hernández, Mireia, María-Ángeles Palomar-García, Benito Nohales-Nieto, Gustau Olcina-Sempere Esteban Villar-Rodríguez, Raúl Pastor, César Ávila and Maria-Antònia Parcet. 2019. “Separate contribution of striatum volume and pitch discrimination to individual differences in music reward”, *Psychological Science* 30, 9: 1352-1361.

Higgins, John. 2019. Minimal pairs for English RP. (http://l2phon.lt.cityu.edu.hk/marlodge/min-pairs/minimal.html) (date of access: 1 Jun. 2023).

Intartaglia, Bastien, Travis White-Schwoch, Nina Kraus and Daniele Schön. 2017. “Music training enhances the automatic neural processing of foreign speech sounds”, *Scientific Reports* 7: 12631.

Jackendoff, Ray. 2009. “Parallels and non-parallels between language and music”, *Music Perception* 26: 195-204.

Jassem, Wiktor. 2003. “Polish”, *Journal of the International Phonetic Association* 33, 1: 103-107.

Jekiel, Mateusz. 2022. “L2 rhythm production and musical rhythm perception in advanced learners of English”, *Poznań Studies in Contemporary Linguistics* 58, 2: 315-340.

Jekiel, Mateusz and Kamil Malarski. 2021. “Musical hearing and musical experience in second language English vowel acquisition”, *Journal of Speech, Language, and Hearing Research* 64, 5: 1666-1682.

Jekiel, Mateusz and Kamil Malarski. 2023. “Musical hearing and the acquisition of foreign-language intonation”, *Studies in Second Language Learning and Teaching* 13, 1: 151-178.

Karma, Kai. 2007. "Musical aptitude definition and measure validation: Ecological validity can endanger the construct validity of musical aptitude tests", *Psychomusicology: A Journal of Research in Music Cognition* 19, 2: 79-90.

Kaszycka, Mariola. 2021. "A study on the relationship between musical ability and EFL pronunciation proficiency", *Linguistics Beyond and Within* 7: 58-67.

Kato, Misaki and Melissa Michaud Baese-Berk. 2020. "The effect of input prompts on the relationship between perception and production of non-native sounds", *Journal of Phonetics* 79: 100964.

Kempe Vera, Dennis Bublitz and Patricia J. Brooks. 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.

Kendall, Tyler and Erik R. Thomas. 2010. Vowels: Vowel manipulation, normalization, and plotting in R. R package. (cran.r-project.org/web/packages/vowels/index.html) (date of access: 1 Jun. 2023).

Kraus, Nina and Bharath Chandrasekaran. 2010. "Music training for the development of auditory skills", *Nature Reviews Neuroscience* 11: 599-605.

Kraus, Nina, Erika Skoe, Alexandra Parbery-Clark and Richard Ashley. 2009. "Experience-induced malleability in neural encoding of pitch, timbre, and timing: Implications for language and music", *Annals of the New York Academy of Sciences* 1169, 1: 543-557.

Langus, Alan, Jacques Mehler and Marina Nespor. 2017. "Rhythm in language acquisition", *Neuroscience and Biobehavioral Reviews* 81: 158-166.

Law, Lily N. C. and Marcel Zentner. 2012. "Assessing musical abilities objectively: Construction and validation of the profile of music perception skills", *PLOS ONE* 7: e52508.

Lemhöfer, Kristin and Mirjam Broersma. 2012. "Introducing LexTALE: A quick and valid lexical test for advanced learners of English", *Behavior Research Methods* 44, 2: 325-343.

Levis, John M. 2018. *Intelligibility, oral communication, and the teaching of pronunciation*. Cambridge: Cambridge University Press.

Lewicka-Mroczek, Ewa and Dorota Szymaniuk. 2013. "Prosody learning strategies and what English philology students know about them", *Crossroads. A Journal of English Studies* 3: 22-34.

Lidji, Pascale, Pierre Jolicoeur, Régine Kolinsky, Patricia Moreau, John F Connolly and Isabelle Peretz. 2010. "Early integration of vowel and pitch processing: A mismatch negativity study", *Clinical Neurophysiology* 121: 533-541.

Llanes-Coromina, Judith, Pilar Prieto and Patrick L. Rohrer. 2018. "Brief training with rhythmic beat gestures helps L2 pronunciation in a reading aloud task", in: Katarzyna Klessa, Jolanta Bachan, Agnieszka Wagner, Maciej Karpiński and Daniel Śledziński (eds.), *Proceedings of the 9th International Conference on Speech Prosody 2018.* Poznań: International Speech Communication Association, 498-502.

Lorenzen, Jane. 2019. "Tonal recall: Musical ability and toneme recognition", in: John Levis, Charles Nagle and Erin F. Todey (eds.), *Proceedings of the 10th Pronunciation in Second Language Learning and Teaching Conference 2019.* Ames: Iowa State University, 307-317.

Loui, Psyche, Kenneth Kroog, Jennifer Zuk, Ellen Winner and Gottfried Schlaug. 2011. "Relating pitch awareness to phonemic awareness in children: Implications for tone-deafness and dyslexia", *Frontiers in Psychology* 2: 111.

Lund, Rebecca M. 2023. "Musical aptitude and the acquisition of pitch and duration in L2 vowels", *Leviathan: Interdisciplinary Journal in English* 9.

Magne, Cyrille, Deanna K. Jordan and Reyna L. Gordon. 2016. "Speech rhythm sensitivity and musical aptitude: ERPs and individual differences", *Brain and Language* 153-154: 13-19.

Magne, Cyrille, Daniele Schön and Mireille Besson. 2006. "Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches", *Journal of Cognitive Neuroscience* 18, 2: 199-211.

Mairano, Paolo and Antonio Romano. 2010. "Un confronto tra diverse metriche ritmiche usando Correlatore", in: Stephan Schmid, Michael Schwarzenbach and Dieter Studer-Joho (eds.), *La dimensione temporale del parlato.* Torriana: EDK, 79-100.

Mairano, Paolo and Antonio Romano. 2011. "Rhythm metrics for 21 languages", in: Wai-Sum Lee and Eric Zee (eds.), *Proceedings of the 17th International Congress of Phonetic Sciences 2011*. Hong Kong: City University of Hong Kong, 1318-1321.

Mandell, Jacob C. 2009. Music tests. (http://jakemandell.com) (date of access: 1 Jun. 2023).

Mandell, Jacob C., Katrin Schulze and Gottfried Schlaug. 2007. "Congenital amusia: An auditory-motor feedback disorder?", *Restorative Neurology and Neuroscience* 25(3-4): 323-334.

Marie, Céline, Franco Delogu, Giulia Lampis, Marta Olivetti Belardinelli and Mireille Besson. 2010. "Influence of musical expertise on segmental and tonal processing in Mandarin Chinese", *Journal of Cognitive Neuroscience* 23: 2701-2715.

Marques, Carlos, Sylvain Moreno, São Luís Castro and Mireille Besson. 2007. "Musicians detect pitch violation in a foreign language better than nonmusicians: Behavioral and electrophysiological evidence", *Journal of Cognitive Neuroscience* 19, 9: 1453-1463.

Masataka, Nobuo. 2007. "Music, evolution and language", *Developmental Science* 10, 1: 35-39.

McAuliffe, Michael, Michaela Socolof, Sarah Mihuc, Michael Wagner and Morgan Sonderegger. 2017. "Montreal forced aligner: Trainable text-speech alignment using Kaldi", in: Francisco Lacerda (ed.), *Proceedings of Interspeech 2017, 18th Annual Conference of the International Speech Communication Association*. Stockholm: International Speech Communication Association, 498-502.

Mennen, Ineke and Esther de Leeuw. 2014. "Beyond segments: Prosody in SLA", *Studies in Second Language Acquisition* 36: 183-194.

Mikoś, Michael J. 1976. "Intonation of questions in Polish", *Journal of Phonetics* 4: 247-253.

Milovanov, Riia, Minna Huotilainen, Paulo A. A. Esquef, Paavo Alku, Vesa Välimäki and Mari Tervaniemi. 2009. "The role of musical aptitude and language skills in preattentive duration processing in school-aged children", *Neuroscience Letters* 460, 2: 161-165.

Milovanov, Riia, Minna Huotilainen, Vesa Välimäki, Paulo A. A. Esquef and Mari Tervaniemi. 2008. "Musical aptitude and second language pronunciation skills in school-aged children: Neural and behavioral evidence", *Brain Research* 1194: 81-89.

Milovanov, Riia, Päivi Pietilä, Mari Tervaniemi and Paulo A. A. Esquef. 2010. "Foreign language pronunciation skills and musical aptitude: A study of Finnish adults with higher education", *Learning and Individual Differences* 20, 1: 56-60.

Mithen, Steven. 2005. *The singing Neanderthals: The origins of music, language, mind and body*. London: Weidenfeld and Nicolson.

Moritz, Catherine, Sasha Yampolsky, Georgios Papadelis, Jennifer Thomson and Maryanne Wolf. 2013. "Links between early rhythm skills, musical training, and phonological awareness", *Reading and Writing: An Interdisciplinary Journal* 26, 5: 739-769.

Nardo, Davide and Susanne M. Reiterer. 2009. "Musicality and phonetic language aptitude", in: Grzegorz Dogil and Susanne M. Reiterer (eds.), *Language talent and brain activity.* New York: De Gruyter Mouton: 213-255.

Nettl, Bruno. 2000. "An ethnomusicologist contemplates universals in musical sound and musical culture", in: Nils L. Wallin, Merker Björn and Steven Brown (eds.), *The origins of music*. Cambridge: MIT Press, 463-472.

Ning, Li Hsin. 2020. "Musical memory and pitch discrimination abilities as correlates of vocal pitch control for speakers with different tone and musical experiences", in: Nobuaki Minematsu, Mariko Kondo, Takayuki Arai and Ryoko Hayashi (eds.), *Proceedings of the 10th International Conference on Speech Prosody 2020.* Tokyo: International Speech Communication Association, 611-615.

Nisha, Kavassery Venkateswaran, Devi Neelamegarajan, Nishant N. Nayagam, Jim Saroj Winston and Sam Publius Anil. 2021. "Musical aptitude as a variable in the assessment of working memory and selective attention tasks", *Journal of Audiology and Otology* 25, 4: 178-188.

Nowacka, Marta. 2012. "Questionnaire-based pronunciation studies: Italian, Spanish and Polish students' views on their English pronunciation", *Research in Language* 10, 1: 43-61.

Nowacka, Marta. 2016. "English spelling among the top priorities in pronunciation teaching: Polglish local versus global(ised) errors in the production and perception of words commonly mispronounced", *Research in Language* 14, 2: 123-148.

Nowacka, Marta. 2022. "English phonetics course: University students' preferences and expectations", *Research in Language* 20, 1: 71-84.

Ordin, Mikhail and Leona Polyanskaya. 2015. “Perception of speech rhythm in second language: The case of rhythmically similar L1 and L2”, *Frontiers in Psychology* 6: 316.

Ott, Cyrill G. M. and Lutz Jäncke. 2013. “Processing of self-initiated speech-sounds is different in musicians”, *Frontiers in Human Neuroscience* 7: 1-15.

Palomar-García, María-Ángeles, Gustau Olcina-Sempere, Mireia Hernández, María-Antonia Parcet, Jacob C. Mandell, César Ávila. 2020. “The Jake Mandell test as a measure of individual differences in pitch discrimination: Validity and reliability properties”, *Revista Electrónica Complutense de Investigación en Educación Musical* 17: 143-151.

Parbery-Clark, Alexandra, Adam Tierney, Dana L. Strait and Nina Kraus. 2012. “Musicians have fine-tuned neural distinction of speech syllables”, *Neuroscience* 219: 111-119.

Pastuszek-Lipińska, Barbara. 2008. “Musicians outperform nonmusicians in speech imitation”, *Lecture Notes in Computer Science* 4969: 56-73.

Patel, Aniruddh D. 2008. *Music, language and the brain*. New York: Oxford University Press.

Patel, Aniruddh D., Jessica M. Foxton and Timothy D. Griffiths. 2005. “Musically tone-deaf individuals have difficulty discriminating intonation contours extracted from speech”, *Brain and Cognition* 59: 310-313.

Pickering, Lucy and Caroline Wiltshire. 2000. “Pitch accent in Indian-English teaching discourse”, *World Englishes* 19, 2: 173-183.

Porzuczek, Andrzej. 2007. “English vowel quantity in Polish learner’s speech perception and production”, in: Janusz Arabski, Danuta Gabryś-Barker and Andrzej Łyda (eds.), *PASE papers 2007. Studies in language and methodology of teaching foreign languages*. Katowice: Para, 96-105.

Reddy, Sravana and James Stanford. 2015. “DARLA: A web application for automated dialect analysis”, in: Matt Gerber, Catherine Havasi and Finley Lacatusu (eds.), *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations*. Denver: Association for Computational Linguistics, 71-75.

Richter, Karin. 2021. “Foreign accent and the role of identity in the adult English as a foreign language pronunciation classroom”, in: Armin Berger, Helen Heaney, Pia Resnik, Angelika Rieder-Bünemann and Galina Savukova (eds.), *Developing advanced English language competence. English language education.* Vol. 22. Springer, Cham, 323-341.

Rojczyk, Arkadiusz. 2010. "Forming new vowel categories in second language speech: The case of Polish learners' production of English /ɪ/ and /e/", *Research in Language* 8: 85-97.

Rojczyk, Arkadiusz. 2011. "Overreliance on duration in nonnative vowel production and perception: The within lax vowel category contrast", in: Magdalena Wrembel, Małgorzata Kul and Katarzyna Dziubalska-Kolaczyk (eds.), *Achievements and perspectives in SLA of speech: New Sounds 2010*. Vol. 2. Frankfurt am Main: Peter Lang.

Rojczyk, Arkadiusz and Andrzej Porzuczek. 2012. "Selected aspects in the acquisition of English phonology by Polish learners – Segments and prosody", in: Danuta Gabryś-Barker (ed.), *Readings in second language acquisition.* Katowice: Uniwersytet Śląski, 93-120.

Rojczyk Arkadiusz and Andrzej Porzuczek. 2017. "Discrimination of English tone contours by Polish learners", *Linguistica Silesiana* 38: 53-66.

Roncaglia-Denissen, Paula M., Drikus A. Roor, Ao Chen and Makiko Sadakata. 2016. "The enhanced musical rhythmic perception in second language learners", *Frontiers in Human Neuroscience* 10: 288

Rosenfelder, Ingrid, Josef Fruehwald, Keelan Evanini, Scott Seyfarth, Kyle Gorman, Hilary Prichard and Jiahong Yuan. 2014. FAVE (forced alignment and vowel extraction) program suite v1.2.2 10.5281/zenodo.22281 (date of access: 1 Jun. 2023).

Rousseau, Jean-Jacques. 1781. *Essai sur l'origine des langues*. Paris: Flammarion.

Saito, Kazuya. 2015. "The role of age of acquisition in late second language oral proficiency attainment", *Studies in Second Language Acquisition* 37, 4: 713-743.

Saito, Kazuya. 2021. "What characterizes comprehensible and native-like pronunciation among English-as-a-second-language speakers? Meta-analyses of phonological, rater, and instructional factors", *TESOL Journal* 55: 866-900.

Sakai, Mari and Colleen Moorman. 2018. "Can perception training improve the production of second language phonemes? A meta-analytic review of 25 years of perception training research", *Applied Psycholinguistics* 39: 187-224.

Schmitt, Holger. 2007. “The case for the epsilon symbol (ɛ) in RP dress”, *Journal of the International Phonetic Association* 37: 321-328.

Schön, Daniele, Cyrille Magne and Mireille Besson. 2004. “The music of speech: Music training facilitates pitch processing in both music and language”, *Psychophysiology* 41: 341-349.

Schwartz, Geoffrey. 2010. “Rhythm and vowel quality in accents of English”, *Research in Language* 8: 135-147.

Schwartz, Geoffrey. 2019. “Voice quality and L2 proficiency in the English tense-lax contrast”, *Anglophonia* [Online], 27. (http://journals.openedition.org/anglophonia/2058) (date of access: 1 Jun. 2023).

Schwartz, Geoffrey and Jerzy Dzierla. 2018. “Polish listeners’ perception of vowel inherent spectral change in L2 English”, *Poznań Studies in Contemporary Linguistics* 54, 2: 307-332.

Schwartz, Geoffrey and Kamil Kaźmierski. 2020. “Vowel dynamics in the acquisition of L2 English – an acoustic study of L1 Polish learners”, *Language Acquisition* 27, 3: 227-254.

Seashore, Carl E., Don Lewis and Joseph G. Saetveit. 1960. *Seashore measures of musical talents*. New York: Psychological Corporation.

Setter, Jane. 2008. “Communicative patterns of intonation in L2 English teaching and learning: The impact of discourse approaches”, in: Katarzyna Dziubalska-Kołaczyk and Joanna Przedlacka (eds.), *English pronunciation models: A changing scene*. Bern: Peter Lang, 367-392.

Slevc, L. Robert. 2012. “Language and music: sound, structure, and meaning.” *Wiley interdisciplinary reviews. Cognitive science* 3, 4: 483-492.

Slevc, L. Robert and Akira Miyake. 2006. “Individual differences in second language proficiency: Does musical ability matter?”, *Psychological Science* 17: 675-681.

Smit, Ute. 2002. “The interaction of motivation and achievement in advanced EFL pronunciation learners”, *International Review of Applied Linguistics in Language Teaching* 40, 2: 89-116.

Sobkowiak, Włodzimierz. 2008. *English phonetics for Poles*. (3rd edition.) Poznań: Wydawnictwo Poznańskie.

Strait, Dana L., Alexandra Parbery-Clark, Emily Hittner and Nina Kraus. 2012. "Musical training during early childhood enhances the neural encoding of speech in noise", *Brain and Language* 123: 191-201.

Strauss, Hannah, Stephan Reiche, Maximilian Dick and Marcel Zentner. 2023. "Online assessment of musical ability in 10 minutes: Development and validation of the Micro-PROMS", *Behavior Research Methods*, Advance online publication. https://doi.org/10.3758/s13428-023-02130-4

Suzukida, Yui. 2021. "The contribution of individual differences to L2 pronunciation learning: Insights from research and pedagogical implications", *RELC Journal* 52, 1: 48-61.

Swaminathan, Swathi and Glenn Schellenberg. 2020. "Musical ability, music training, and language ability in childhood", *Journal of Experimental Psychology: Learning, Memory, and Cognition* 46, 12: 2340-2348.

Szyszka, Magdalena. 2018. "Pronunciation learning environment: EFL students' cognitions of in-class and out-of-class factors affecting pronunciation acquisition", *Theory and Practice of Second Language Acquisition* 4, 1: 121-139.

Turker, Sabrina and Susanne M. Reiterer. 2021. "Brain, musicality, and language aptitude: A complex interplay", *Annual Review of Applied Linguistics* 41: 95-107.

Wallentin, Mikkel, Andreas Højlund Nielsen, Morten Friis-Olivarius, Christian Vuust and Peter Vuust. 2010. "The musical ear test: A new reliable test for measuring musical competence", *Learning and Individual Differences* 20, 3: 188-196.

Wang, Hao, Peggy Mok and Helen Meng. 2016. "Capitalizing on musical rhythm for prosodic training in computer-aided language learning", *Computer Speech and Language* 37, C: 67-81.

Wang, Xianhui, Chao-Yang Lee and Seth Wiener. 2023. "Non-native disadvantage in spoken word recognition is due to lexical knowledge and not type/level of noise", *Speech Communication* 149: 29-37.

Waniek-Klimczak, Ewa, Arkadiusz Rojczyk and Andrzej Porzuczek. 2015. "'Polglish' in Polish eyes: What English studies majors think about their pronunciation in English", in: Ewa Waniek-Klimczak and Mirosław Pawlak (eds), *Teaching and researching the pronunciation of English. Second language learning and teaching.* Springer, Cham, 23-23.

Weckwerth, Jarosław. 2011. “English TRAP vowel in advanced Polish learners: Variation and system typology”, in: Wai-Sum Lee and Eric Zee (eds.), *Proceedings of the 17th International Congress of Phonetic Sciences 2011*. Hong Kong: City University of Hong Kong, 2110-2113.

Weinberger, Steven. 2015. Speech accent archive. George Mason University. (http://accent.gmu.edu) (date of access: 1 Jun. 2023).

Wells, J. C. 1982. *Accents of English*. Cambridge: Cambridge University Press.

Wells, J. C. 2014. *English intonation: An introduction*. Cambridge: Cambridge University Press.

White, Donald and Peggy Mok. 2018. “L2 speech rhythm development in new immigrants”, in: Katarzyna Klessa, Jolanta Bachan, Agnieszka Wagner, Maciej Karpiński and Daniel Śledziński (eds.), *Proceedings of the 9th International Conference on Speech Prosody 2018.* Poznań: International Speech Communication Association, 838-842.

White, Laurence and Sven L. Mattys. 2007. “Rhythmic typology and variation in first and second languages”, in: Pilar Prieto, Joan Mascaró and Maria-Josep Solé (eds.), *Segmental and prosodic issues in Romance phonology. Current issues in linguistic theory series.* Amsterdam: John Benjamins, 237-257.

Willems, Nico. 1982. *English intonation from a Dutch point of view*. Dordrecht: Foris Publications.

Wong, Patrick C. M. and Tyler K. Perrachione. 2007. “Learning pitch patterns in lexical identification by native English-speaking adults”, *Applied Psycholinguistics* 28, 4: 565-585.

Wong, Patrick C. M., Erika Skoe, Nicole M. Russo, Tasha Dees and Nina Kraus. 2007. “Musical experience shapes human brainstem encoding of linguistic pitch patterns”, *Nature Neuroscience* 10, 4: 420-422.

Zatorre, Robert J. and Shari R. Baum. 2012. “Musical melody and speech intonation: Singing a different tune”, *PLoS Biology* 10, 7: e1001372.

Zhang, J. Diana, Marco Susino, Gary E. McPherson and Emery Schubert. 2020. “The definition of a musician in music psychology: A literature review and the six-year rule”, *Psychology of Music* 48, 3: 389-409.

Zybert, Jerzy and Sabina Stępień. 2009. “Musical intelligence and foreign language learning”, *Research in Language* 7: 99-111.

Appendix A: Research article 1 (Jekiel and Malarski 2021): Author contribution statement (Polish)

Poznań, 01.06.2023

Oświadczenie o wkładzie autorów

Jekiel, Mateusz, and Kamil Malarski. 2021. "Musical hearing and musical experience in second language English vowel acquisition", *Journal of Speech, Language, and Hearing Research 64*, 5: 1666-1682. https://doi.org/10.1044/2021_JSLHR-19-00253

Autor oraz współautor artykułu naukowego „Musical hearing and musical experience in second language English vowel acquisition", opublikowanego w *Journal of Speech, Language, and Hearing Research* 11 maja 2021 r., niniejszym deklarują swój następujący wkład w powstanie artykułu:

- **Mateusz Jekiel** (60%): projekt badania (konceptualizacja), przygotowanie materiałów, zbieranie danych, analiza danych, wizualizacja danych, napisanie manuskryptu (przygotowanie oryginalnego tekstu, recenzja i redagowanie);

- **Kamil Malarski** (40%): projekt badania (konceptualizacja), zbieranie danych, napisanie manuskryptu (przygotowanie oryginalnego tekstu, recenzja i redagowanie).

AUTORZY:

Mateusz Jekiel
....
mgr Mateusz Jekiel

Kamil Malarski
....
dr Kamil Malarski

PROMOTOR PRACY DOKTORSKIEJ:

Piotr G...
....
dr hab. Piotr Gąsiorowski, prof. UAM

Appendix B: Research article 1 (Jekiel and Malarski 2021): Author contribution statement (English)

Poznań, 01.06.2023

Author contribution statement

Jekiel, Mateusz, and Kamil Malarski. 2021. "Musical hearing and musical experience in second language English vowel acquisition", *Journal of Speech, Language, and Hearing Research 64*, 5: 1666-1682. https://doi.org/10.1044/2021_JSLHR-19-00253

Both co-authors of the research article „Musical hearing and musical experience in second language English vowel acquisition", published in *Journal of Speech, Language, and Hearing Research* on 11th May 2021, hereby declare that they have contributed to the research article in the following way:

- **Mateusz Jekiel** (60%): study design (conceptualisation), material preparation, data collection, data analyses, data visualisation, manuscript writing (original draft preparation, reviewing, and editing);

- **Kamil Malarski** (40%): study design (conceptualisation), data collection, manuscript writing (original draft preparation, reviewing, and editing).

AUTHORS:

Mateusz Jekiel
....

mgr Mateusz Jekiel

Kamil Malarski
....

dr Kamil Malarski

PhD SUPERVISOR:

Piotr Gąsiorowski
....

dr hab. Piotr Gąsiorowski, prof. UAM

Appendix C: Research article 2 (Jekiel 2022): Author contribution statement (Polish)

Poznań, 01.06.2023

Oświadczenie o wkładzie autora

Jekiel, Mateusz. 2022. "L2 rhythm production and musical rhythm perception in advanced learners of English", *Poznan Studies in Contemporary Linguistics 58*, 2: 315-340. https://doi.org/10.1515/psicl-2022-0016

Autor artykułu naukowego „L2 rhythm production and musical rhythm perception in advanced learners of English", opublikowanego w *Poznan Studies in Contemporary Linguistics* 17 czerwca 2022 r., niniejszym deklaruje swój następujący wkład w powstanie artykułu:

- **Mateusz Jekiel** (100%): projekt badania (konceptualizacja), przygotowanie materiałów, zbieranie danych, analiza danych, wizualizacja danych, napisanie manuskryptu (przygotowanie oryginalnego tekstu, recenzja i redagowanie).

AUTOR:

Mateusz Jekiel

....

mgr Mateusz Jekiel

PROMOTOR PRACY DOKTORSKIEJ:

Piotr Gąs

....

dr hab. Piotr Gąsiorowski, prof. UAM

Appendix D: Research article 2 (Jekiel 2022): Author contribution statement (English)

Poznań, 01.06.2023

Author contribution statement

Jekiel, Mateusz. 2022. "L2 rhythm production and musical rhythm perception in advanced learners of English", *Poznan Studies in Contemporary Linguistics 58*, 2: 315-340. https://doi.org/10.1515/psicl-2022-0016

The author of the research article "L2 rhythm production and musical rhythm perception in advanced learners of English", published in *Poznan Studies in Contemporary Linguistics* on 17th June 2022, hereby declare that they have contributed to the research article in the following way:

- **Mateusz Jekiel** (100%): study design (conceptualisation), material preparation, data collection, data analyses, data visualisation, manuscript writing (original draft preparation, reviewing, and editing).

AUTHOR:

Mateusz Jekiel
....

mgr Mateusz Jekiel

PhD SUPERVISOR:

Piotr Gąs
....

dr hab. Piotr Gąsiorowski, prof. UAM

Appendix E: Research article 3 (Jekiel and Malarski 2023): Author contribution statement (Polish)

Poznań, 01.06.2023

Oświadczenie o wkładzie autorów

Jekiel, Mateusz, and Kamil Malarski. 2023. "Musical hearing and the acquisition of foreign-language intonation", *Studies in Second Language Learning and Teaching 13*, 1: 151-178. https://doi.org/10.14746/ssllt.23166

Autor oraz współautor artykułu naukowego „Musical hearing and the acquisition of foreign-language intonation", opublikowanego w *Studies in Second Language Learning and Teaching* 31 marca 2023 r., niniejszym deklarują swój następujący wkład w powstanie artykułu:

- **Mateusz Jekiel** (60%): projekt badania (konceptualizacja), przygotowanie materiałów, zbieranie danych, analiza danych, wizualizacja danych, napisanie manuskryptu (przygotowanie oryginalnego tekstu, recenzja i redagowanie);

- **Kamil Malarski** (40%): projekt badania (konceptualizacja), zbieranie danych, napisanie manuskryptu (przygotowanie oryginalnego tekstu, recenzja i redagowanie).

AUTORZY:

Mateusz Jekiel
....
mgr Mateusz Jekiel

Kamil Malarski
....
dr Kamil Malarski

PROMOTOR PRACY DOKTORSKIEJ:

Piotr Gąsi
....
dr hab. Piotr Gąsiorowski, prof. UAM

Appendix F: Research article 3 (Jekiel and Malarski 2023): Author contribution statement (English)

Poznań, 01.06.2023

Author contribution statement

Jekiel, Mateusz, and Kamil Malarski. 2023. "Musical hearing and the acquisition of foreign-language intonation", *Studies in Second Language Learning and Teaching 13*, 1: 151-178. https://doi.org/10.14746/ssllt.23166

Both co-authors of the research article "Musical hearing and the acquisition of foreign-language intonation", published in *Studies in Second Language Learning and Teaching* on 31st March 2023, hereby declare that they have contributed to the research article in the following way:

- **Mateusz Jekiel** (60%): study design (conceptualisation), material preparation, data collection, data analyses, data visualisation, manuscript writing (original draft preparation, reviewing, and editing);

- **Kamil Malarski** (40%): study design (conceptualisation), data collection, manuscript writing (original draft preparation, reviewing, and editing).

AUTHORS:

Mateusz Jekiel
....

mgr Mateusz Jekiel

Kamil Malarski
....

dr Kamil Malarski

PhD SUPERVISOR:

....

dr hab. Piotr Gąsiorowski, prof. UAM

Appendix G: Research article 1 (Jekiel and Malarski 2021)

Jekiel, Mateusz and Kamil Malarski. 2021. "Musical hearing and musical experience in second language English vowel acquisition", *Journal of Speech, Language, and Hearing Research 64*, 5: 1666-1682. https://doi.org/10.1044/2021_JSLHR-19-00253

JSLHR

Research Article

Musical Hearing and Musical Experience in Second Language English Vowel Acquisition

Mateusz Jekiel[a] and Kamil Malarski[a]

Purpose: Former studies suggested that music perception can help produce certain accentual features in the first and second language (L2), such as intonational contours. What was missing in many of these studies was the identification of the exact relationship between specific music perception skills and the production of different accentual features in a foreign language. Our aim was to verify whether empirically tested musical hearing skills can be related to the acquisition of English vowels by learners of English as an L2 before and after a formal accent training course.
Method: Fifty adult Polish speakers of L2 English were tested before and after a two-semester accent training in order to observe the effect of musical hearing on the acquisition of English vowels. Their L2 English vowel formant contours produced in consonant–vowel–consonant context were compared with the target General British vowels produced by their pronunciation teachers. We juxtaposed these results with their musical hearing test scores and self-reported musical experience to observe a possible relationship between successful L2 vowel acquisition and musical aptitude.
Results: Preexisting rhythmic memory was reported as a significant predictor before training, while musical experience was reported as a significant factor in the production of more native-like L2 vowels after training. We also observed that not all vowels were equally acquired or affected by musical hearing or musical experience. The strongest estimate we observed was the closeness to model before training, suggesting that learners who already managed to acquire some features of a native-like accent were also more successful after training.
Conclusions: Our results are revealing in two aspects. First, the learners' former proficiency in L2 pronunciation is the most robust predictor in acquiring a native-like accent. Second, there is a potential relationship between rhythmic memory and L2 vowel acquisition before training, as well as years of musical experience after training, suggesting that specific musical skills and music practice can be an asset in learning a foreign language accent.

The assumed deep-binding interconnectivity between musical talent and the ease of learning foreign languages has been anecdotally repeated for a long time, yet too often without clear references to the relevant research. Indeed, there are many similarities between language and music, and both have been thoroughly studied over the years. Studies in music and language evolution argue that humans developed musicality in similar ways they developed their language capacities (Brown, 2001; Mithen, 2005; see also Baudeat, 2017). Another large portion of research focused on the similarities of how music and language are processed on the neural level (Brown et al., 2006; Chobert & Besson, 2013; Fadiga et al., 2009; Kunert et al., 2015). Considering these commonalities, language researchers and trainers have looked into the possible effects that music perception and production might have on the acquisition of various language structures (Pastuszek-Lipińska, 2008; Strait et al., 2012), as well as on its role in language teaching (Fonseca-Mora et al., 2011; Franklin et al., 2008; Picavet et al., 2012). Following this line of research, the current study is interested whether preexisting musical skills and musical experience can predict subsequent learning outcomes in the acquisition of second language (L2) pronunciation in a formal learning environment.

Despite the large body of research concerning the relationship between musical hearing and L2 learning, there is still considerable room for exploration due to the limitations of previous studies. First, many prior experiments relied on self-reported language proficiency (Roncaglia-

[a]Faculty of English, Adam Mickiewicz University, Poznań, Poland
Correspondence to Mateusz Jekiel: mjekiel@amu.edu.pl
Editor-in-Chief: Bharath Chandrasekaran
Editor: Chao-Yang Lee
Received September 29, 2019
Revision received December 23, 2019
Accepted January 4, 2021
https://doi.org/10.1044/2021_JSLHR-19-00253

Disclosure: *The authors have declared that no competing interests existed at the time of publication.*

Denissen et al., 2016), limited the notion of L2 proficiency to grammar and vocabulary (e.g., Franklin et al., 2008: musical training – verbal memory) or resorted to perceptual evaluation of L2 pronunciation (Milovanov et al., 2010). As far as music skills are concerned, many studies divided participants into musicians and nonmusicians based on their music education or assessed their musical aptitude via formal music education tests, which relied on both production and perception (Ott et al., 2011).

The primary aim of our research is to study the relationship between different aspects of musical hearing and the acquisition of L2 vowel inventory by Polish advanced learners of English. We conducted three musical hearing tests assessing pitch perception, melodic memory, and rhythmic memory, while music education and musical experience were treated independently. In order to assess segmental features of L2 pronunciation, we performed vowel measurements of our participants before and after a two-semester accent training course at a university level and compared their results with the model vowels of their pronunciation teachers. From all the second-language accent features, we have selected vowels because they can be reliably measured acoustically, while their signal properties are most similar to the music sequences comprising different tonalities. To summarize, the twofold goal of this study is to determine whether musical skills and musical experience can predict successful acquisition of L2 vowels during an accent training course, as well as to verify if these preexisting musical factors can be positively associated with L2 pronunciation prior to formal training.

Music, Language, and the Brain

Psycholinguists and cognitive neuroscientists (e.g., Chobert & Besson, 2013; Patel, 2008) emphasize that music and speech share similarities on various levels: (a) They are both specific to human beings and exist across all known cultures; (b) both are complex auditory signals that have tone, melody, and rhythm; (c) both can be described in terms of frequency, duration, intensity, and timbre; (d) both have several degrees of organization (morphology, phonology, semantics, syntax, and pragmatics in language; rhythm, melody, and harmony in music); (e) both systems require attention, memory, and sensorimotor coordination for accurate perception and production; and (f) both share neural resources for processing prosody, syntax, and semantics.

Neuroimaging and neuroscientific studies comparing the brain responses of musicians and nonmusicians when listening to musical and linguistic stimuli found functional and structural differences between both groups, suggesting that the properties of the transverse temporal gyri in the brain can be associated with musical aptitude (e.g., Schneider et al., 2002; see Barrett et al., 2013). Studies using functional magnetic resonance imaging and positron emission tomography also show an overlap in specific brain regions responsible for processing of both linguistic and musical structures, predominantly the Broca's area (e.g., Brown et al., 2006; Fadiga et al., 2009; Kunert et al., 2015), as well as greater brain plasticity and higher functional restructuring of the brain across people who had musical training at an earlier age (Gaser & Schlaug, 2003; Hyde et al., 2009); it has been shown that even short-termed musical training can be conducive to such neural changes (Lappe et al., 2008). While the abovementioned studies reveal an interesting relationship between musical skills and linguistic skills, Schellenberg (2004) and Patel (2012) point out that the majority of behavioral and neural data in favor of this hypothesis are not derived from experimental studies that would allow us to draw strong causal inferences. Importantly, however, a longitudinal study was conducted by Moreno et al. (2009) to directly test for causality (i.e., causal effects of music vs. painting training on different aspects of speech perception). Using both behavioral and electrophysiological methods, their results revealed that musical training causally impacts on linguistic abilities of children (i.e., positive transfers from musical training to speech processing). However, it still remains unclear which specific aspects of musical hearing potentially influence phonemic processing skills dependent on other acoustic properties. For example, if the speaker's timbre, an auditory attribute that relies on a different acoustic dimension than pitch, can be affected by musical training focused on pitch perception, it would suggest that both domains "share more abstract cognitive processes involved in sound categorization, possibly involving common cortical mechanisms" (Patel, 2012, p. 29).

The Role of Vowels in Language and Music

The sound systems in language and music are multimodal, complex, and therefore difficult to describe. Consonants can display secondary articulations; while producing vowels, in turn, speakers of different languages may make use of different tones that can change the meaning of an utterance. In music, timbre is also important, which becomes apparent when the same musical piece is transcribed to be played on a different instrument. In sung pieces, however, language and music are integrated: The singing is treated like an instrument, capable of operating on a particular timbre and pitch. It has been argued before that, at an earlier stage of language development, the modulations in timbre and pitch were more common (Nikolsky, 2015), and similar connections between vowel production and musical pitch are supported in experimental research on the processing of vowels and melody (Kolinsky et al., 2009; Russo et al., 2019). It has also been concluded that interacting neural networks are responsible for processing differences in phonemes and pitch in singing, treating them as integrated units (Lidji et al., 2010).

Vowels can be described as more music-like than consonants. First, Fenk-Oczlon (2017) emphasizes the importance of vowels for both language and music, pointing out their role in generating sonority in syllables and prosody in speech and singing. Similarly to musical notes, vowels have timbre, pitch, intensity, and duration. Moreover, the number of vowels often corresponds to the number of pitches in musical scales across cultures: Just like most languages have a vowel inventory comprising five vowels (Maddieson, 2013), pentatonic scales (i.e., musical scales with five notes per octave) are also considered to be the most common among traditional forms of music (Trehub et al., 1999). Finally, the musical quality of vowels plays a significant role in early language acquisition, especially in the use of pitch and timbre in early communication for pragmatic functions (Masataka, 2007). Due to these commonalities in the acoustic and auditory nature of vowels and music, we have selected this phonemic category for our analyses.

The common procedure in linguistics has been to present vowels in a two-dimensional space, most commonly on vowel charts, to approximate the position of the tongue during their production. Both the height/openness (F1) and frontness/backness (F2) of vowels are expressed in Hertz (Hz). In our analysis, we will focus only on the first (F1) and the second (F2) vowel formant characteristics because we are only interested in how high/low and front/back they are produced in our participants. Although it is possible for a person to pronounce a given vowel on different fundamental frequencies, the human brain must have a way to decode the F1 and F2 frequencies to hear the given vowel quality (see, e.g., Ikeda et al., 2014; Tankus et al., 2012).

Perception and Production

The relationship between perception and production in L2 pronunciation has been widely investigated over the years, frequently revolving around the notion of causal relationship between these two domains. The Speech Learning Model (Flege, 1995, 2007) postulates that the accuracy of production of nonnative sounds is correlated with their perception. This assumes that improved performance in perception is required for improved performance in production. Several studies have both confirmed and rejected this hypothesis (see Isbell, 2016). Another approach to study the relation between perception and production is correlational—recent research revealed moderate correlation between L2 perception and L2 production of segmental contrasts (e.g., Casillas, 2015); however, some studies found no correlation between these domains (cf. Zhang et al., 2016).

Latest findings show that production can shape perception, which means a language learner has to be constantly exposed to the spoken real-life target language structures; otherwise, production may disrupt the perceptual learning skills (Baese-Berk, 2019). An individual's first language (L1) can also influence their auditory mechanisms (e.g., phonetic properties of L1 speech can predict the performance in perception of L2 vowels across speakers; Kartushina & Frauenfelder, 2014). Moreover, speakers of different language can also process music and music-like sounds differently, for example, Chinese speakers demonstrate superior pitch processing (Bidelman et al., 2011), while Finnish speakers display more precise discrimination of duration (Dawson et al., 2017). Such differences are associated with the structure of an L1 and its cognitive representation in the speaker.

Musical Hearing and L2 Acquisition

Successful acquisition of L2 pronunciation relies on a number of linguistic and extralinguistic factors. The former include similarities and differences between L1 and L2 phonetic inventories, while the latter involve such aspects as the onset age of exposure to L2, motivation, learning strategies, working memory, and musical experience. According to Chobert and Besson (2013, p. 923), musical hearing "positively influences several aspects of speech processing, from auditory perception to speech production." In this study, we are interested whether musical hearing is associated with the production of L2 vowels before and after formal accent training.

According to Dolman and Spring (2014), musical hearing can be defined as an untaught, natural musical ability, and should be regarded as separate from musical experience, which includes music lessons and musicianship. One traditional method of assessing musical hearing is the Seashore Measures Of Musical Talents (Seashore et al., 1960), a standardized test that divides musical hearing into separate talents (focusing on pitch, duration, rhythm, timbre and tonality, loudness, etc.). However, this type of assessment is very formal and used primarily at music schools, where the participants are potential candidates already familiar with the fundamentals of music theory and practice. Here, we use a series of musical hearing tests devised by Mandell (2009), which rely on music perception and can be attempted by participants without any musical knowledge or formal instruction.

Recent studies suggest that musical hearing and training can influence second language acquisition, especially the acquisition of L2 pronunciation, including segmental and suprasegmental vocalic discrimination (Chobert & Besson 2013). Musical aptitude seems to correlate with phonological processing ability among preschool children (e.g., Anvari et al., 2002), while musical training has been associated with better linguistic skills (Strait et al., 2012; Tallal & Gaab, 2006), preexisting musical abilities seem to be better predictors for the acquisition of language skills than formal training in music (Swaminathan & Schellenberg, 2019). In a study by Slevc and Miyake (2006), superior musical hearing correlated with more accurate L2 listening discrimination and production skills among Japanese adult learners of English, while Milovanov et al. (2010) observed a relationship between musical skills and pronunciation skills among Finnish adult learners of English producing challenging English phonemes in a speech shadowing task. These findings motivated the current research, which is aimed at establishing whether musical hearing skills and former musical experience can be a valid predictor for successful acquisition of L2 pronunciation among Polish adult learners of English.

Polish Learners of English

The Polish vowel system is much less complicated than the one we find in English. Also, dialectal variation in vowels is marginal, as opposed to the large differentiation in English.

Polish has six vowels (/a/, /ɛ/, /i/, /ɨ/, /ɔ/, /u/) that do not differ significantly in length phonemically. Considering the rich vowel inventory in English, where vowels differ both in terms of quantity and quality, it is natural that Polish students of English usually find it difficult to master these contrasts. One problematic vowel contrast is the KIT–FLEECE distinction, in terms of both production and perception (Rojczyk & Porzuczek, 2012, p. 99; Sobkowiak, 2008), despite the fact that, in Polish, there are seemingly comparable vowel categories /i/ and /ɨ/. The TRAP vowel also poses many problems for Polish learners, even at advanced levels of language proficiency—the reason for this is the lack of a counterpart for this vowel in Polish, resulting in the assimilation to /ɛ/ or /a/ (Weckwerth, 2011).

To the best of our knowledge, there have only been a few studies investigating the role of music in the acquisition of L2 English phonology featuring participants with L1 Polish. A study by Pastuszek-Lipińska (2008) showed that Polish learners of English with formal music expertise (i.e., musical education and training) tend to outperform learners without any musical background in imitation and shadowing tasks. In this series of observations, musicians were better at discriminating foreign language sounds and produced fewer errors. A similar study by Zybert and Stępień (2009) utilized tests assessing music perception (Edwin Gordon's *Intermediate Measure of Music Audiation*) and production (imitating musical intervals), along with a speech perception and production test involving textbook audio recordings of a native speaker of English. The results for Polish secondary school learners of English with and without music education showed a correlation between perception and production of language and music, suggesting that both music perception and production can be good predictors in successful L2 learning. A fairly recent pilot study by Gralińska-Brawata and Rybińska (2017) investigated the production of word stress and its connection with musical abilities, where Polish advanced learners of English were asked to answer a questionnaire related to music, perform a music perception and production test, and read a set of commonly mispronounced words. The results suggest a possible relation between the correct word stress and superior musical abilities, although the authors pointed out the need for a larger sample. Since most of the abovementioned studies relied on imitation, shadowing, and repetition, the following study can be regarded as a relevant supplement to this discussion, as it incorporates visual stimuli and relies on learners' phonemic awareness developed throughout an intensive pronunciation training course.

This Study

The objective of the study is to investigate whether specific aspects of musical hearing and musical experience can predict successful acquisition of L2 vowels by Polish learners of English before and after formal accent training. The influence of musical hearing and musical training on the development of language skills, especially in the contexts of language acquisition and pedagogy, has recently received significant attention. The studies reported above, however, do not answer many of the questions we were interested in investigating. First, shadowing or imitation tasks do not explain the entire process in which L2 sounds are acquired, and in-class imitation is often dissimilar from the actual long-term formation of the L2 vowel inventory. For this reason, we conducted recording sessions before and after training, which relied on reading aloud from a screen a list of words without providing an immediate pronunciation model, which should offer a better insight into the acquisition of L2 pronunciation. Second, in order to sufficiently account for the musical hearing of our participants, we performed three different musical hearing tests assessing different aspects of musical hearing. We were interested in whether specific aspects of musical hearing, such as pitch perception, melodic memory, and rhythmic memory, are translatable into phonetic and phonological skills in our participants. Finally, we also included music education and musical experience in our analyses as separate factors that can relate to L2 vowel acquisition. To summarize, we formulate the following research questions:

1. Is there an observable progress in participants' L2 vowel production after training? Does the participants' distribution of vowel formants become more similar to their pronunciation teachers' vowel formants after training?
2. Are participants with better musical skills better at learning to pronounce L2 vowels?
3. Does musical experience predict how well participants learn to pronounce L2 vowels?

Method

Participants

Our subjects were 50 native speakers of Polish (42 women, eight men). Their age varied between 19 and 21 years (M = 20.14, SD = 0.40). At the time of the study, they were enrolled in Year 1 of a 3-year bachelor's program (1BA) in English studies, which included modules in linguistics, literature, culture, and EFL (English as a Foreign Language). All participants were recruited from four different groups that followed the same curriculum. At the time of the recording, their level of English was between B2 and C1 within the Common European Framework of Reference framework (a guide for categorizing European learners of foreign languages in terms of their achievement; Council of Europe, 2011). It was a homogenous group in this respect as they had performed very similarly in their secondary school final exams in written and oral English (Polish *matura* exams). To confirm their language proficiency, we also conducted LexTALE (Lexical Test for Advanced Learners of English) by Lemhöfer and Broersma (2012), commonly used to study participants with an advanced level of L2 English in an experimental setting. The results confirmed that the group was uniform and fairly advanced (M = 74.48%, SD = 8.93). Despite their advanced command of written and spoken English, none of our participants had obtained substantial or regular pronunciation instruction at an earlier stage, which is not an exception, since pronunciation training in the EFL classroom is not systematically taught in primary or secondary education (Derwing & Munro, 2015). The first recording session was scheduled for the first week of their university education in order to factor out the influence of the accent training they were to receive during the academic year. Thus, so far, they had learnt their EFL pronunciation solely through naive exposure. None of our subjects had any medically documented speech or hearing impairments. They took part in the experiment in exchange for extra course credit. All participants were volunteers and were not financially remunerated for their participation. Informed written consent was obtained from all participants prior to their inclusion in the study.

Accent Training

All participants were assigned to a two-semester accent training course and attended the classes twice a week during the first year of their studies (a total of 90 hr of class work). The course comprised segmental phonetics (i. e., vowels and consonants of English) and suprasegmental phonetics (i.e., syllable stress, sentence stress, intonation, and rhythm), and its primary goal was to allow the participants to develop a consistent, native-like pronunciation based on the General British (henceforth GB) accent (i.e., the accent commonly used in education and usually associated with the South of England[1]). The course syllabus included such topics as English vowels (monophthongs and diphthongs) and consonants (fortis–lenis distinction and its effect on the preceding vowel), connected speech processes (e.g., assimilation, elision, linking), word stress, sentence stress, and weak forms.

Four different pronunciation teachers conducted the classes. The teachers were all female native speakers of Polish with native-like GB accents and were experienced language trainers with knowledge in phonetics and English language teaching. As the teachers shared the same L1 with the learners, they could rely on their individual experience in mastering English pronunciation and help the learners to avoid errors considered as typical for Polish learners of English (e.g., spelling pronunciation, difference in the tongue position, inconsistent vowel duration or lack of vowel reduction; see Sobkowiak, 2008). We recorded the teachers using the same procedure as for our participants in order to compare their vowel formant contours. The formant frequencies for their L2 vowels were similar to the model GB values found in Cruttenden (2014, p. 104): We observed a strong correlation between the teacher's L2 F1/F2 vowel formants and the model formant values ($r = .99$, $df = 19$, $p < .001$). Phonetic instruction used in the accent training course was holistic (i.e., instead of teaching the sounds in isolation, coarticulation, connected speech processes, and stress were incorporated throughout the course). In the classroom, the primary teaching methods involved exposure, phonetic transcription, production via drills and reading exercises, improvisation, and prepared speeches. Outside the classroom, all participants had to deliver monthly recordings based on the practiced reading material. Assessment relied on mid-semester and end-semester exams, which included a word list, a reading text, and spontaneous speech. The resulting accent was to be native-like and devoid of marked L1 features. After the training, all participants were expected to acquire native-like GB pronunciation through successful recognition and production of phonemic contrasts and native-like allophony.

English Phonetics and Phonology Course

The accent training course is also supplemented by an obligatory two-semester course in English phonetics and phonology, which provides a phonological description of the English sound system and raises phonological awareness. The primary aim of the course is to introduce students to the field of descriptive linguistics and help them as future educators in correcting errors made by Polish learners of English. The course was taught once a week for a total of 45 hr of class work, and the syllabus included the following topics: articulatory phonetics, acoustic phonetics, phonemic and phonetic transcription, English and Polish phonemes and allophones, connected speech processes, and prosody. Student assessment was based on weekly online quizzes and mid- and end-semester tests.

Recordings

Two recording sessions before and after training took place in a recording studio in a sound-treated booth. A single session, including the interview and musical hearing tests, took roughly 30 min for each participant. Each recording session was structured as a sociolinguistic interview and started with a casual conversation to minimize stress and help participants adapt to the experimental setting. Afterward, the main part of the recording session started, which comprised a series of word lists in English to elicit all GB monophthongs in the most formal and controlled context. The word list featured vowels in various consonant–vowel–consonant frames (b_d, b_t, d_d, d_t, g_d, h_d, and h_t) with five repetitions per word. The obtained vowels were 10 monophthongs: DRESS, KIT, FOOT, LOT, STRUT, TRAP, FLEECE, GOOSE, THOUGHT, and START. We obtained a total of 50 tokens in the h_d frame for this study for comparability with similar studies on vowel production, which often rely on this consonantal context. The word list was divided into 10 sections, each beginning with similarly phrased instructions: "These words rhyme with ___ or ___," (e.g., "These words rhyme with *feed* or *feet*"). The instructions were visualized on the screen and read by the investigator. Next, the participants were presented with the tokens belonging to that category on the monitor screen, always one

[1]At the beginning of their studies, students can choose to attend General British or General American accent training.

word per screen. The words were displayed in a large sans-serif font, in white letters against a black background. Each word was presented for 2,000 ms, with a 300-ms prerecording delay between each token. During that time frame, participants had to read aloud the token word that appeared on the screen. The stimulus was prepared and presented with SpeechRecorder software (Draxler & Jänsch, 2019), operated from a MacBook Pro. To capture the voice signal, we used an MXL 770 microphone, connected to a Roland Duo Capture EX audio interface. We were able to communicate with the participants sitting in the booth through AKG K240 MKII studio headphones and a PreSonus talkback monitor station. The speech samples were recorded at a mono 44.1-kHz frequency and 16-bit resolution, and were saved as separate wave files.

Musical Hearing Tests

After the first recording session but before accent training, each participant participated in three musical hearing tests designed by Mandell (2009), which were tested on over 11,000 subjects. Each test focuses on one musical hearing skill: pitch perception, melodic memory, and rhythmic memory, features that are regarded as valid indicators of musical aptitude and that are similarly tested in other musical hearing tests (e.g., Wallentin et al., 2010). The tests were run on a laptop connected to studio headphones.

The pitch perception test (Adaptive Pitch Test[2]) is designed to measure pitch perception abilities by playing a series of two short tones and asking if the second tone is higher or lower than the first. The participants used the UP and DOWN arrows on the keyboard to choose if the second tone was higher or lower. They could also replay the tones using the spacebar. The duration of the pitch perception test was approximately 1 min. Next, in the melodic memory test (Tonedeaf Test), each participant had to determine whether 36 pairs of short instrumental melodies were the same or different from one another by pressing the corresponding button on the screen. No repetition was possible in this test. The test was designed to verify pitch perception and melodic memory, as well as identify neuroanatomical correlates of tone deafness and was used as a screening test for patients tested in Mandell et al. (2007). The previously reported results had shown that the test was relatively difficult for both clinical (Mandell et al., 2007) and nonclinical (Mandell, 2009) participants; therefore, it is expected that our participants will also score low in this test. It took about 5 min to complete. Finally, in the rhythmic memory test (Rhythm Test), each participant was asked to decide whether each pair of 25 short rhythmic instrumental patterns was the same or different from each other by pressing the corresponding button on the screen. The test was designed to assess the ability to distinguish subtle differences in rhythm. The duration of this test was circa 5 min.

[2]The names in brackets are the names of the tests developed by Mandell (2009) available online.

Musical Experience Questionnaire

After the first recording session, we asked our participants in a questionnaire whether they attended music school or have specific musical experience (i.e., singing as soloists, band members or choir members, and/or playing a musical instrument as soloists, band members, or members of an orchestra or other ensemble). We also asked to specify the years spent in music school and/or years spent practicing singing and/or playing a musical instrument. Participants who reported no formal musical training or experience and who reported to sing or play a musical instrument only occasionally were treated in the analysis as nonmusicians.

Vowel Measurements

The data were prepared for analysis in Praat software (Boersma & Weenink, 2019), and the measurements were obtained through FAVE-Extract (Rosenfelder et al., 2014) and Montreal Forced Aligner (McAuliffe et al., 2017) using the DARLA web interface (Reddy & Stanford, 2015) and the *Vowels* R package (Kendall & Thomas, 2010). Since our participants were taught by different pronunciation teachers, we normalized their vowel measurement results separately for each group, including their teachers. We selected the Fabricius et al. (2009) method, which is speaker-intrinsic but vowel-extrinsic and formant-extrinsic. This normalization technique calculates a speaker's vowel space on the basis of the closest and the most open vowel in GB, which are often the TRAP and FLEECE vowels, respectively; in the Polish vowel system, the closest vowel is /i/, whereas the most open is /a/. By means of normalized vowel charts with the vowels of our participants plotted, we were able to graphically present the differences between their L2 vowels and the pronunciation model. The more similar their L2 vowels were to the model GB vowels, the better their EFL pronunciation would be evaluated. In order to quantify for the spatial relations between the vowels, we calculated the Euclidean distances between our participants' L2 vowels and teacher's model vowels (comparing F1 and F2 dimensions), where x1 stands for teacher's F1 and x2 for participant's F1, while y1 stands for teacher's F2 and y2 for participant's F2.

$$\text{distance} < \text{-sqrt}\left((x1-x2)^\wedge 2+(y1-y2)^\wedge 2\right). \qquad (1)$$

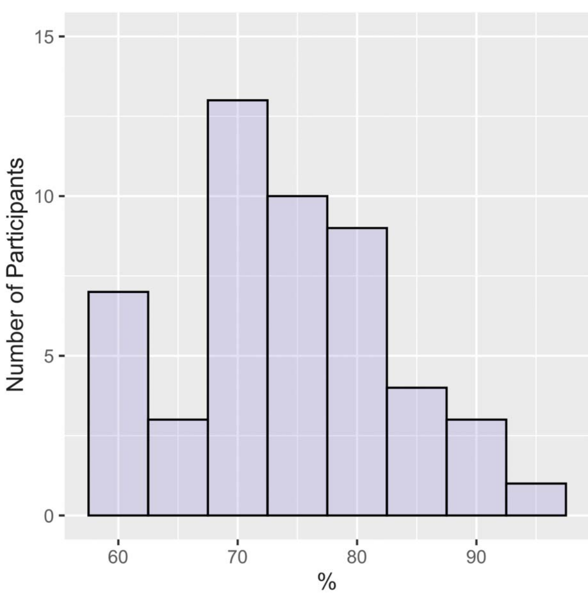
In the Results section, we are discussing the differences between the mean formant values of individual vowels produced by the participants and the corresponding mean vowel formants produced by the teachers. These vowel distance scores were then compared with the musical hearing test results, music education, years of musical experience, LexTALE results, and gender, using linear mixed-effects models.

Results

LexTALE

The LexTALE test results presented in Figure 1 confirmed that the subjects of the study formed a relatively

Figure 1. LexTALE results (%), the higher the better (M = 74.47, SD = 8.92).



uniform group in terms of their EFL proficiency. The horizontal axis is the test score (i.e., the percentage of correct responses). The mean result was 74.47 (min = 60, max = 97.50, *Mdn* = 73.75).

Musical Hearing Tests

The pitch perception test results presented in Figure 2 are expressed in Hertz (Hz) and indicate how reliably the participant could differentiate between two tones. The mean result was 16.05 Hz (min = 1 Hz, max = 60 Hz, *Mdn* = 10.02). The default grading for the test is fairly rigorous: The ability to reliably differentiate two tones less than 0.75 Hz apart indicates an exceptional ear (none of the participants achieved that score), while less than 1.5 Hz equals as fairly good (scored by 2 participants), less than 6 Hz as normal (13 participants), less than 12 Hz as low normal (13 participants) and above 16 Hz as below normal, possibly indicating a pitch perception deficit (18 participants).

The melodic memory test results in Figure 3 are expressed in percentages and indicate the percent of correctly identified tokens. According to the default scoring for this test, a result below 70% indicates low performance (scored by 27 participants), 70%–79% is normal (17 participants), 80%–90% is above normal (six participants), and above 90% is exceptional (none of the participants achieved that score). The mean result was 68% (min = 44.4%, max = 86.1%, *Mdn* = 69.4%).

The rhythmic memory test results presented in Figure 4 are expressed in percentages and indicate the percent of correctly identified tokens. Similarly to the previous test, a result below 70% indicates low performance, 70%–79% is normal, 80%–90% is above normal, and above 90% is exceptional. The mean result was 71.2% (min = 48%, max = 92%, *Mdn* = 72%).

Music Education and Musical Experience

Out of 50 participants, 18 confirmed to have had some musical experience. Four participants confirmed they had attended and completed the first degree of music school —two of them have played a musical instrument for 12 years

Figure 2. Pitch perception test results (Hz), the lower the better (M = 11.80, SD = 15.86).

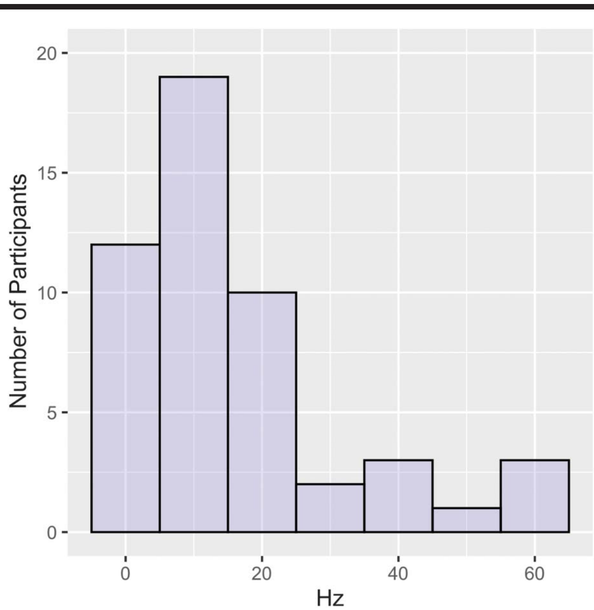


Figure 3. Melodic memory test results (%), the higher the better (M = 67.90, SD = 9.58).

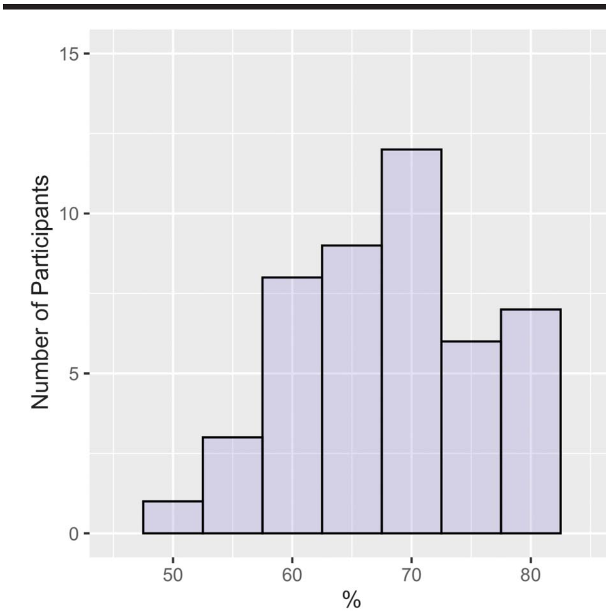


Figure 4. Rhythmic memory test results (%), the higher the better (M = 71.20, SD = 10.35).

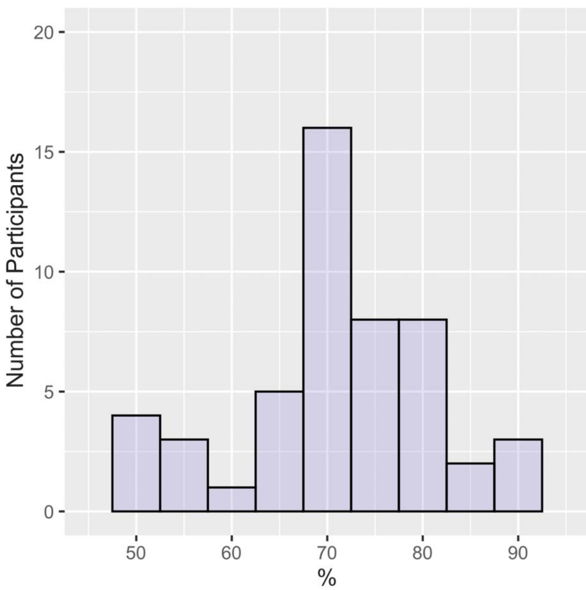
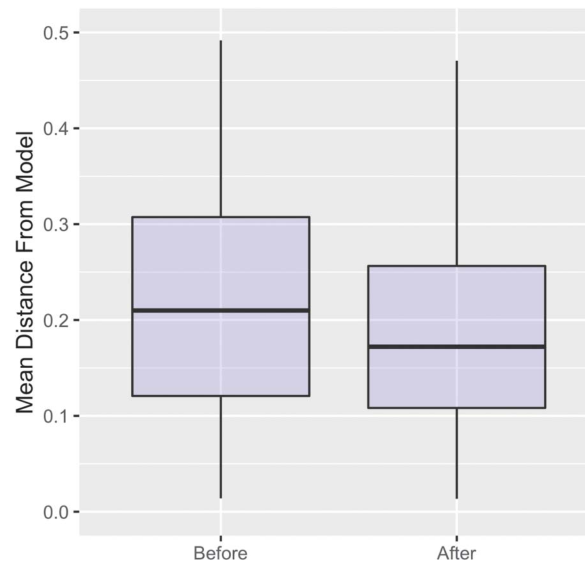


Figure 5. Distance from GB model before (M = 0.23, SD = 0.13) and after training (M = 0.20, SD = 0.12). The difference between the results is significant at $t(499) = -5.46$, $p < .001$.



and the other two have played for 6 years. Eight other participants reported to have played a musical instrument without formal music education—three of them have played for 9 years, four have played for 6 years, and one has played for 3 years. Thirteen participants confirmed singing either as soloists or band/choir members. Two of them were the same participants who had completed music school, while five were the same who reported to have played a musical instrument without formal training. Six other participants have sung for 12 years but have not played a musical instrument nor received music education. The summarized data can be found in Table 1. For the linear mixed-effects models, we treated music education as a categorical variable and musical experience as a numerical variable with years of musical practice of our participants as instrumentalists and/or vocalists (whichever was higher).

Euclidean Distance From GB Model Before and After Training

Figure 5 shows the mean distance from the GB model for all vowels before and after training for all 50 participants. A parametric t test for dependent means confirmed a significant difference between mean distance from the GB model before training (M = 0.23, SD = 0.13) and after training (M = 0.20, SD = 0.12), $t(499) = -5.46$, $p < .001$.

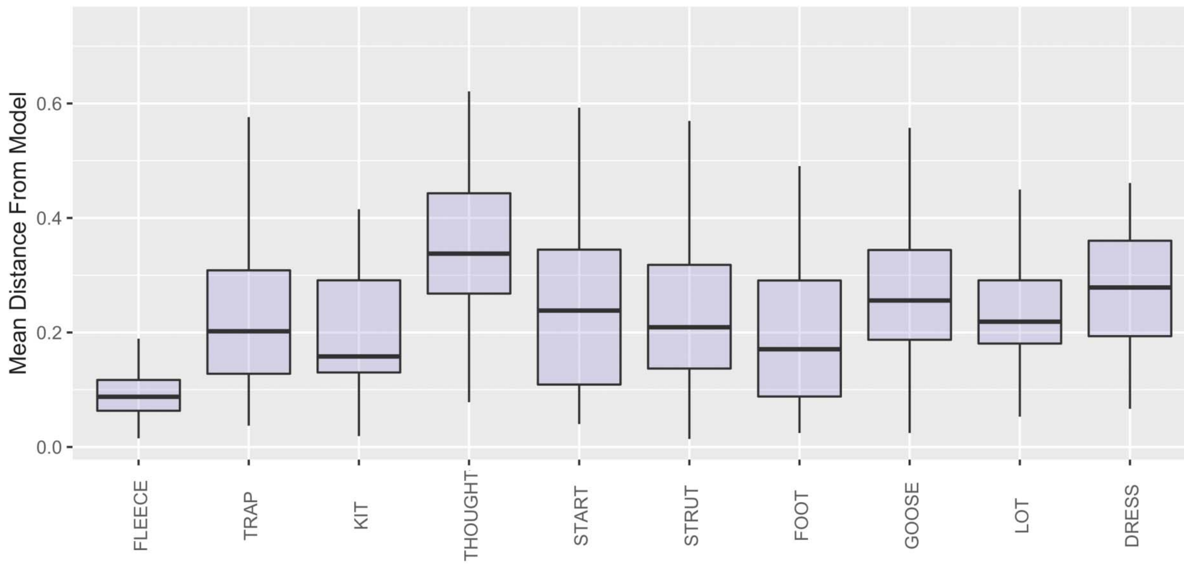
Table 1. Number of participants with self-reported musical experience.

Years	Music education	Musical instrument	Singing
12	2	2	10
9	—	3	—
6	2	6	2
3	—	1	1

Figure 6 shows the mean distance from the GB model for each individual vowel before training. A one-way analysis of variance reported a significant difference between the average measures of the 10 vowels, $F(9, 490) = 14.45$, $p < .01$, suggesting that the participants had already found some vowels to be more problematic in production than others. FLEECE produced by the participants was the closest vowel to the GB model (M = 0.09, SD = 0.05), while the most distant was THOUGHT (M = 0.34, SD = 0.13).

Figure 7 shows the mean distance from the GB model for each individual vowel after training. A one-way analysis of variance reported a similarly significant difference between the average measures of the 10 vowels, $F(9, 490) = 15.16$, $p < .01$, indicating that not all L2 vowels were equally mastered by the participants after training. Again, FLEECE produced by the participants was the closest vowel to the GB model and had little variation (M = 0.08, SD = 0.04). Next, the vowels with similar variation and mean distance from the model were TRAP, KIT, THOUGHT, START, and STRUT. More variation could be observed for back vowels (i.e., FOOT, LOT, and GOOSE, the latter exhibiting more significant variation [M = 0.28, SD = 0.15]). It is worth noting at this point that both FOOT and GOOSE vowels are often confused by Polish learners of English, as there is only one /u/ sound in their L1 vowel system. Furthermore, while Polish has a vowel similar to LOT in its phonetic inventory (i.e., the open-mid back /ɔ/, many Polish learners of English often substitute it with the General American /ɑː/), mostly due to their exposure to

Figure 6. Distance from General British model before training for individual 10 monophthongs.



American English at earlier stages of education and the predominant American English accents present in popular culture. Interestingly, the vowel, which, on average was the most extremely distant from the model, was DRESS (M = 0.27, SD = 0.12), although it should be a fairly easy vowel to acquire by Polish learners of English because, in their native vowel inventory, there exists a nearly equivalent open-mid front vowel /ɛ/. It would seem then that developing a new L2 vowel category would only entail raising the Polish /ɛ/ vowel slightly.

Figure 8 shows two contrasting participants producing GB vowels before training. P45 (left) managed to produce vowel sounds that were similar to the GB model, while the tokens of P29 (right) were further from the model. It is apparent that, while P45 has a more consistent and categorical distribution of vowel sounds, the same tokens for P29 often overlap or show a considerable degree of variation. Figure 9 shows the same two participants after their 1-year accent training. It can be observed that the former acquired a more categorical distribution of vowels with a relatively low degree of variation, while the latter still has apparent overlaps in their vowel system and a higher degree of variation for some vowels.

Figure 7. Distance from General British model after training for individual 10 monophthongs.

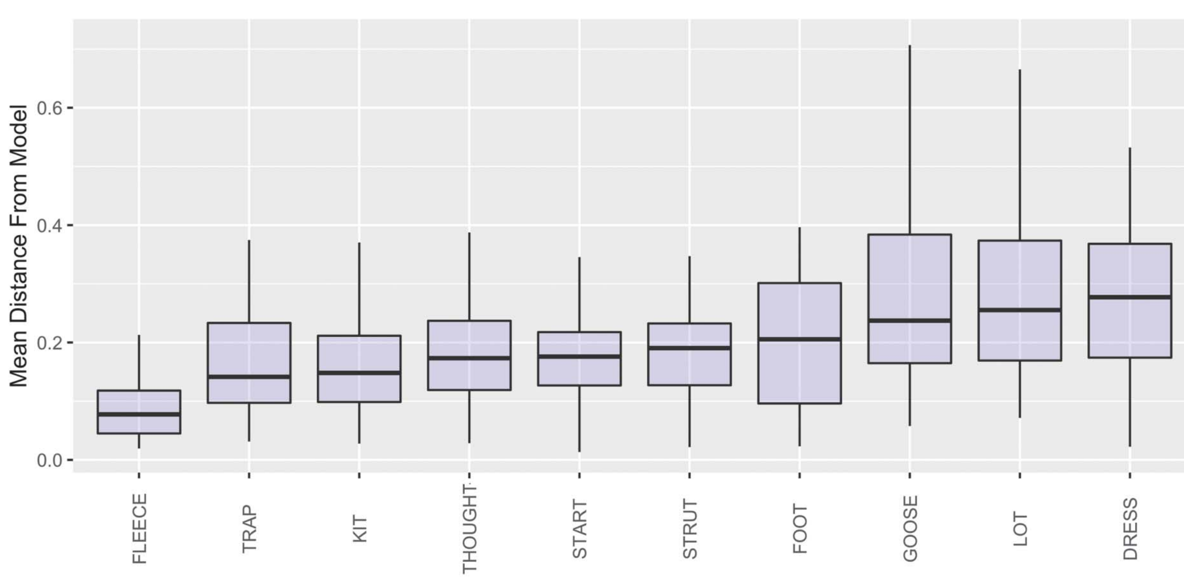
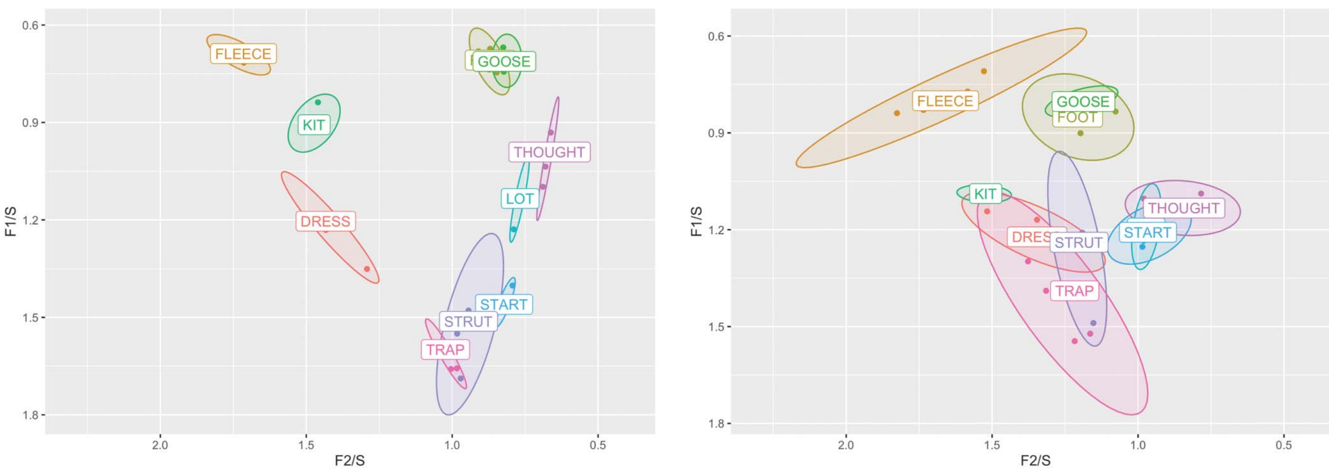


Figure 8. Example vowel formants before training for P45 (left) and P29 (right).



Closeness to Model Before Training

Prior to investigating participants' vowel production after training, we looked into the effects of individual variables on their results before the accent training course. A linear mixed-effects regression model was built to explain the closeness of vowel formants produced by the participants' to the model values before training, with speaker and vowel as random effects and gender, LexTALE score, pitch perception test result, melodic memory test result, rhythmic memory test result, music education, and years of musical experience as fixed effects. The results for 500 observations of 50 speakers and 10 vowels presented in Table 2 show a significant result for rhythmic memory (p = .004). A negative estimate (−.002) for that parameter indicates that a more accurate rhythmic memory is associated with a smaller distance from the pronunciation model. Therefore, participants who scored higher on that particular musical hearing test (i.e., were better at differentiating two rhythmic patterns) were also closer to the model. For example, participant P016 with the highest score (92%) in the rhythmic memory test was also the closest to the pronunciation model, with .164 in their mean closeness to the model before training. Conversely, participant P029 with one of the lowest results in this test (52%) was the furthest from the model before training (.374). We found no significant results for the pitch perception test results or the melodic memory test results. Likewise, we also found no relation between participants' closeness to model before training and their music education or musical experience. We reported no multicollinearity between the independent variables in the VIF scores (< 1.5).

For the estimated random effects, we found a significant difference between variance for speaker (.0001, SD = 0.01) and vowel groups (.0039, SD = 0.06), which suggests that the differences between vowels can be more informative

Figure 9. Example vowel formants after training for P45 (left) and P29 (right).

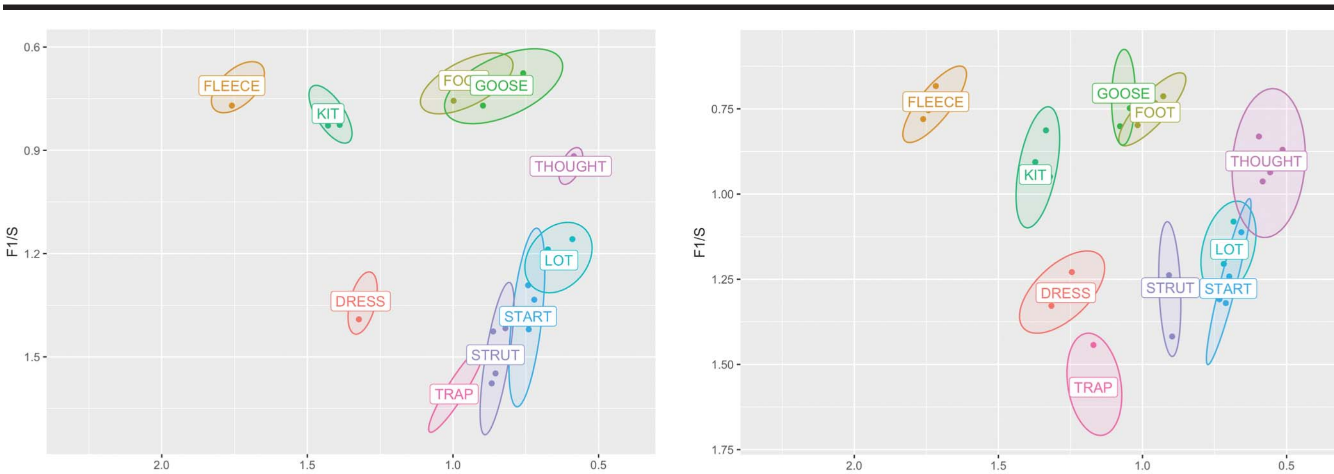


Table 2. Linear mixed-effects model results for closeness to model before training (500 observations, 50 speakers, 10 vowels).

Parameter	Estimate	*SE*	Test (df)	*p*
Intercept	0.290	0.07	4.14 (47.75)	< .001
Gender (M)	−0.006	0.02	−0.34 (41.99)	.733
LexTALE	0.000	0.00	0.30 (41.99)	.762
Pitch perception	0.000	0.00	1.28 (41.99)	.207
Melodic memory	0.000	0.00	1.10 (41.99)	.275
Rhythmic memory	**−0.002**	**0.00**	**−3.01 (41.99)**	**.004**
Music education	0.001	0.02	0.07 (41.99)	.945
Musical experience	0.000	0.00	0.38 (41.99)	.702

Note. *SE* = standard error; M = male; LexTALE = Lexical Test for Advanced Learners of English. Boldface indicates statistical significance ($p \leq .05$).

Table 3. Linear mixed-effects model results for closeness to model after training (500 observations, 50 speakers, 10 vowels).

Parameter	Estimate	*SE*	Test (df)	*p*
Intercept	0.186	0.07	2.82 (48.77)	.006
Gender (M)	0.018	0.02	1.12 (42.01)	.267
LexTALE	−0.000	0.00	−0.67 (42.00)	.506
Pitch perception	−0.000	0.00	−0.47 (42.20)	.641
Melodic memory	−0.000	0.00	−0.49 (42.15)	.628
Rhythmic memory	−0.000	0.00	−0.24 (43.16)	.810
Music education	0.022	0.02	1.02 (41.99)	.312
Musical experience	**−0.002**	**0.00**	**−2.12 (42.01)**	**.039**
Before training	**0.368**	**0.04**	**9.68 (489.1)**	**< .001**

Note. *SE* = standard error; M = male; LexTALE = Lexical Test for Advanced Learners of English. Boldface indicates statistical significance ($p \leq .05$).

than between speakers. We calculated separate linear regressions for each vowel to investigate the relation between rhythmic memory and participants' closeness to model before training. We reported a significant result with a negative estimate −.004 for TRAP, $F(1, 48) = 6.06$, $p = .017$, $R^2 = .09$, and a weak significant result with a negative estimate −.003 for FOOT, $F(1, 48) = 3.63$, $p = .062$, $R^2 = .05$. This is an interesting finding, since Polish students of English often struggle with the pronunciation of TRAP and substitute it with Polish /ɛ/ or /a/ (Weckwerth, 2011), while FOOT is often confused with GOOSE and perceived as Polish /u/ (Balas, 2018). Both vowels were also similarly challenging for our participants before training.

Closeness to Model After Training

To predict the closeness of vowel formants produced by the participants' to the model values after training, we built a linear mixed-effects regression model with speaker and vowel as random effects and gender, LexTALE score, pitch perception test result, melodic memory test result, rhythmic memory test result, music education, years of musical experience, and the closeness to the model before training as fixed effects. The results for 500 observations of 50 speakers and 10 vowels presented in Table 3 show a significant result with a negative estimate for musical experience (−.002, $p = .039$) and a very significant result with a positive estimate for the closeness to model before training (.368, $p < .001$). We reported no multicollinearity between the independent variables in the VIF scores (< 1.5). These results indicate that participants with more years of musical experience also achieved a more native-like pronunciation by producing similar vowels to their pronunciation teachers. For instance, participant P055 with the closest mean values after training (.125) also had 12 years of singing experience, while participant P022 who was the furthest from the model (.307) had no musical experience. We found no significant results for the musical hearing tests, possibly suggesting that either musical hearing is not an asset during explicit pronunciation training in a formal learning environment, or that the training itself is effectively flattening any potential influence of musical hearing on participants' progress. The results also suggest that years of practicing music, either by playing a musical instrument or singing, can be more informative than having formal music education, as we found no significant result for that parameter. However, it is important to point out that we only had four participants who reported attending music school. To further investigate this dependence, it would be required to have a considerable number of participants divided into two equal groups of formally trained musicians and amateur musicians, as well as an insight into their actual musical performance and motivation in their music education or practice.

Similarly to the previous model, we found a significant difference between variance for speaker (.0003, $SD = 0.02$) and vowel (.0024, $SD = 0.05$). We calculated separate linear regressions for each vowel to investigate the relation between musical experience and participants' closeness to model after training. We reported a significant result with a negative estimate −.008 for TRAP, $F(1, 48) = 6.23$, $p = .016$, $R^2 = .09$, and a weak significant result with a negative estimate −.006 for STRUT, $F(1, 48) = 4.37$, $p = .041$, $R^2 = .06$. As previously established, TRAP is a difficult vowel for Polish learners of English, and it is possible that musical practice can affect participants' progress in the acquisition of this particular vowel, while STRUT is also often substituted by Polish /a/ or mispronounced due to spelling pronunciation (Weckwerth, 2011).

Being able to produce vowels similar to the GB model prior to the actual training was a determining factor, confirming that participants who were relatively accurate in their pronunciation from the start were also closer to the model after training. A participant's mean distance from the GB model is expected to increase by .368 for every unit of distance from that model before training ($p < .001$) after controlling for the other variables. While there were participants who produced vowels similar to their pronunciation teachers both before and after training (e.g., P028 before = .192, after = .191), there were also participants who made a significant progress during training (e.g., P029 before = .374, after = .197). However, to determine the effect of musical hearing on the actual progress in the acquisition of L2 pronunciation, it would be required to have a study group comprising participants with fairly similar vowel qualities at the beginning of the study. This is why we decided not to measure the difference between the closeness to model before and after training as such, but rather focus on the participants' performance after training, including their closeness to model at the beginning of the accent training course as one of the explanatory variables.

We also calculated separate linear regressions for each vowel to investigate the relation between participants' closeness to model before and after training. We reported strong significant results with a positive estimate .856 for DRESS, $F(1, 48) = 59.93$, $p < .001$, $R^2 = .55$; .778 for GOOSE, $F(1, 48) = 34.1$, $p < .001$, $R^2 = .40$; .666 for FOOT, $F(1, 48) = 35.06$, $p < .001$, $R^2 = .41$; .311 for FLEECE, $F(1, 48) = 6.85$, $p = .012$, $R^2 = .11$; .262 for THOUGHT, $F(1, 48) = 10.16$, $p = .002$, $R^2 = .16$; and .241 for START, $F(1, 48) = 11.63$, $p = .001$, $R^2 = .18$. Interestingly, most of these vowels are considered as long (i.e., having a dimension not present in the Polish vowel system), while at the same time we found no significant results for KIT, TRAP, STRUT, and LOT (i.e., short vowels, which are often substituted with Polish vowel equivalents by Polish learners of English). This would suggest that learners who produced more native-like long vowels after training had been already familiar with these sounds, while other participants were still far from the model after the pronunciation course.

Finally, the results also show that the LexTALE result had no effect on the expected distance from the GB model, suggesting that a general language proficiency, especially relating to vocabulary, is not related to the acquisition of a native-like accent of English. However, a Pearson *r* test confirmed a positive, though weak correlation between the LexTALE results and the melodic memory ($r = .06$) and rhythmic memory ($r = .11$) tests, suggesting that a follow-up study to investigate the relationship between general language proficiency and specific aspects of musical hearing can be valuable.

Discussion

The study examined the acquisition of L2 vowels by 50 Polish advanced learners of English during a two-semester accent training course and the effect of musical hearing and musical experience on the estimated closeness to the GB pronunciation model. The experiment included an acoustic analysis of vowel formants before and after training, a series of musical hearing tests, and a questionnaire regarding musical experience. According to the results, L2 pronunciation is trainable in adult Polish learners of English, and specific aspects of musical hearing and musical experience can predict successful acquisition of L2 vowels before and after training. Although the learners' former proficiency in L2 pronunciation is the most robust predictor in acquiring a native-like accent, preexisting rhythmic memory is also positively associated with pretraining L2 pronunciation, while years of musical practice can predict better posttraining production of L2 vowels, even when pretraining pronunciation is held constant.

We predicted that the participants should produce L2 vowels closer to the pronunciation model after training. It turned out that there was a significant difference between the participants' vowels before and after training, suggesting that a 1-year pronunciation course can have an effect on L2 vowel acquisition and that L2 pronunciation is teachable in a formal academic context. Moreover, we observed that not all vowels were similarly easy to acquire by our participants: The formant values of front vowels FLEECE, KIT, and TRAP were the closest to the model after training, while the most distanced were back vowels GOOSE, FOOT, and LOT. Interestingly, the DRESS vowel was the most difficult to be acquired by our participants. The reason for this can be twofold. First, the DRESS vowel in English is a near-counterpart of the Polish vowel /ɛ/. Perceptually, they can be quite similar, especially considering the recent trend for the GB vowel DRESS to be produced as lower than in the more conservative normative accent labeled received pronunciation (i.e., British English pronunciation based on educated speech in southern England; cf. Cruttenden, 2014). Therefore, because perceptually both vowels /ɛ/ are so similar, the Polish students did not develop a new category for their English vowel. This is perhaps why the teachers usually do not devote a lot of time for discussing and practicing this vowel. Secondly, what could have added to the effect is that the traditional way of teaching the DRESS vowel in Polish higher education context by some teachers has been to instruct the students to produce the vowel higher than in Polish (i.e., in a more conservative way). When we investigated the formant values in the production of the DRESS vowel by the teachers, it turned out that indeed their DRESS vowel was produced higher than their Polish /ɛ/ vowel, similarly to the received pronunciation values found in Cruttenden (2014). This rather conservative feature of their accents could have been ignored by the students or consciously unlearnt, possibly due to their exposure to other contemporary English accents in the media (i.e., students did not incorporate a conservative language feature to their repertoire, if they had heard it produced differently by the native speakers). There also remains the question to what extent our results are extendable to other language pairs; while Polish learners often have problems with learning the TRAP vowel (Weckwerth, 2011) or the contrasts between FLEECE and KIT vowels (Rojczyk & Porzuczek, 2012), speakers with different L1s may find other contrasts more problematic.

We assumed that participants who received better results in the musical hearing tests should also produce vowels closer to their pronunciation model after training. While the test results assessing pitch perception and melodic memory were not associated with the closeness to the model, we reported that the rhythmic memory test scores were related to the shorter distance between the participants' formant values of vowels and the pronunciation model before training. Although language rhythm is commonly associated with suprasegmental features, studies have also suggested links between rhythm and segmental phonology, particularly relating to the temporal aspects of the tense–lax contrast in English (Schwartz, 2010). Indeed, rhythmic perception can have an important role in successful language acquisition (Jusczyk, 1999) and practicing musical rhythm can help achieve more native-like pronunciation and fluency (Llanes-Coromina et al., 2018). Former studies also confirmed the existence of shared neurocognitive resources for rhythm in music and speech (e.g., Magne et al., 2016), as well as successful use of musical rhythm in teaching L2 English prosody, particularly in practicing duration contrasts between stressed and unstressed syllables (Wang et al. 2016). Moreover, Milovanov et al. (2010) also noted in their study that choir members, who performed better in discriminating rhythmic patterns than nonmusicians, also produced fewer pronunciation errors than nonmusical university students. Since many Polish learners of English struggle with achieving native-like L2 pronunciation due to the complex English vowel system with its durational contrasts and vowel reduction, it is possible that rhythmic memory can help such learners achieve more native-like pronunciation through superior discrimination between tense and lax vowels. However, since our study was interested primarily in vowel quality, it would be valuable to see a follow-up study investigating the role of musical rhythm in the acquisition of L2 English vowels by Polish learners of English in terms of both vowel duration and vowel reduction.

We found no relation between participants' closeness to model before or after training and their pitch perception test results, although it is worth noticing that the reported high values for this test vary from similar previous studies (e.g., Amitay et al., 2006; Micheyl et al., 2006), even though all participants were identically instructed and used the same equipment. Possibly, participants reacted differently to the test due to the relatively long recording session, which preceded the musical hearing test and the weak results might have been caused by fatigue. Alternatively, the fact that the pitch perception test was conducted before the melodic memory test and rhythmic memory test might have affected the results. Finally, this is the first study that used this method on Polish learners of English, so perhaps the weak results stem either from insufficient music education in Polish schools (see Zwolińska, 2008) or from L1 influence on auditory processing, which was not yet tested across Polish learners of English (cf. Dawson et al., 2017). These results suggest that musical hearing is not one concept but that it comprises different abilities, which are also differently applicable to production skills.

We also assumed a potential relation between successful L2 vowel acquisition and musical experience. A significant result for this parameter suggests that participants who spent more years practicing music, either by playing a musical instrument or singing, also produced more native-like vowels after training. At the same time, we found no relationship between music education and successful L2 vowel acquisition before or after training. However, it is important to point out that only four participants reported attending music school, thus requiring further investigation in a more balanced sample. Since this study was conducted among Polish students of English in an academic context, we had no control over their former music education or musical experience. One important parameter that should be considered is the starting age of music education and training, as it can be related with successful language acquisition at an early stage (Brandt et al., 2012), as well as helping young learners with hearing impairment (Torppa & Huotilainen, 2019). This could be a potential point of departure for future studies.

The strongest observable estimate in our study was the closeness to model before training (i.e., participants who managed to have acquired more native-like pronunciation before the two-semester accent training course were also more successful afterwards). Consequently, the effects of the intensive formal instruction in the form of pronunciation training did not factor out all the differences between the participants in this study. Even though the participants were a fairly homogeneous group in terms of their age, LexTALE results measuring their overall language proficiency, as well as studying the same degree, English, which required from them similar secondary final exam results, they varied in terms of their pronunciation before the study. Moreover, all participants confirmed that they had not obtained any pronunciation instruction before the study. This means that, given the differences between the speakers before the training, the learners must have ways of assimilating elements of L2 pronunciation outside the context of formal instruction.

Finally, we observed different relations between the closeness to model for individual vowels and the significant variables from the models. The most interesting finding was for TRAP, a considerably difficult vowel for Polish learners of English who frequently substitute it with /ɛ/ or /a/ (Weckwerth, 2011). While we found no relation between the production of TRAP before and after training, we observed a significant result for rhythmic memory and musical experience, indicating that both musical hearing and musical practice can be strongly associated with the successful acquisition of this particular sound. Rhythmic memory was also found as a significant parameter explaining the closeness to model for FOOT, another difficult vowel, commonly confused with GOOSE, and exhibiting considerable variation among Polish learners of English. For musical experience, we also reported a significant result for STRUT, a problematic vowel that is often mispronounced due to spelling pronunciation. Lastly, we observed that the majority of vowels produced after training with a significant result for the "before training" parameter were long vowels (FLEECE, START, GOOSE, THOUGHT), suggesting that participants who had already acquired these sounds were closer to the model after training, while participants who struggled with the tense–lax contrast before the pronunciation course had still some problems with these vowels. These results suggest that musical hearing and musical experience are not only connected to a more native-like L2 vowel system as a whole, but can be related to specific vowels that pose considerable problems for Polish learners of English.

It is difficult at this point to assess the teaching implications of our results. On the one hand, we were investigating the talent or the aptitude aspect in the process of L2 phonology acquisition because it had been clearly shown in many studies before that motivation and formal instruction both play a significant role in this process. On the other hand, both musical and linguistic perception are trainable. There is a growing body of evidence (Baese-Berk, 2019) suggesting that production can shape perception, precisely in the context of foreign-language learning. It is better to expose learners of English to real-life accent models as early as possible in the classroom setting to help them in successful acquisition of L2 pronunciation (Darcy et al., 2012). Moreover, while this study was primarily focused on the effect of musical hearing on L2 pronunciation, it is also possible that an intensive accent training course could affect musical hearing, as many exercises during the course rely on exposure, including ear training exercises and auditory discrimination tests, which, in turn, can influence the learners' awareness to speech sounds and musical sounds, as well. Furthermore, in recent years, there has been an increasing interest in the importance of learner's motivation over innate talent. While the concept of an inborn gift for pronunciation has been widely researched (e.g., Dogil & Reiterer, 2009; Jilka et al., 2011), the relevance of hard work still remains understudied. Grit, as defined by Duckworth et al. (2007), is the "perseverance and passion for long-term goals," and it can be significant in predicting the learner's success. Therefore, it would be of great value to develop reliable tools for measuring not only learners' talents but also their motivation.

Admittedly, the concept behind this article is based on the comparison of two different domains, which are perception and production. Moreover, the conceptual and statistical model used is multimodal and may not account for all aspects of vowel acquisition in advanced learners of English. Yet speech is a complex, multimodal, and context-dependent means of communication. The proposed model hopefully contributes to a better understanding of the mechanisms involved in the acquisition of L2 vowels and the role of specific aspects of musical hearing and musical experience that play in this process.

Conclusions

This study reported an experiment investigating whether preexisting musical hearing skills and musical experience can predict more native-like production of L2 vowels by Polish advanced learners of English before and after formal accent training. The results show that rhythmic memory is positively associated with more native-like L2 pronunciation before training, while years of musical experience can predict more accurate L2 pronunciation after training. Overall, the mean Euclidean distance between the participants' vowels and the model vowels produced by their pronunciation teachers decreased after training, but the change was more distinct for participants who performed better in the rhythmic memory test or had some musical experience as instrumentalists or vocalists. At the same time, we found no significant relation between the mean distance from the model and music education, suggesting that attending music school on its own might not be a predicting factor in the acquisition of native-like pronunciation. Finally, we found a strong relationship between the results before and after training, indicating that participants' former proficiency was related with their improved performance after the accent training course. Thus, while musical hearing and musical experience can play a role in the acquisition of L2 English vowels, individuals' L2 pronunciation before accent training seems to be the determining factor in explaining the closeness to the model after accent training.

Acknowledgments

This research was supported by the National Science Centre in Poland, Grant 2014/15/N/HS2/03865. Principal investigator: Mateusz Jekiel. Recipient: Adam Mickiewicz University in Poznań, Poland. The authors would like to express their gratitude to Professor Piotr Gąsiorowski for his assistance in the early stages of the project, as well as Professor Robert Lew and Kacper Łodzikowski for their help in the data analysis.

References

Amitay, S., Irwin, A., & Moore, D. (2006). Discrimination learning induced by training with identical stimuli. *Nature Neuroscience, 9,* 1446–1448. https://doi.org/10.1038/nn1787

Anvari, S. H., Trainor, L. J., Woodside, J., & Levy, B. A. (2002). Relations among musical skills, phonological processing, and early reading ability in preschool children. *Journal of Experimental Child Psychology, 83*(2), 111–130. https://doi.org/10.1016/S0022-0965(02)00124-8

Baese-Berk, M. M. (2019). Interactions between speech perception and production during learning of novel phonemic categories. *Attention, Perception, & Psychophysics, 81*(4), 981–1005. https://doi.org/10.3758/s13414-019-01725-4

Balas, A. (2018). English vowel perception by Polish advanced learners of English. *Canadian Journal of Linguistics, 63*(3), 309–338. https://doi.org/10.1017/cnj.2018.5

Barrett, K. C., Ashley, R., Strait, D. L., & Kraus, N. (2013). Art and science: How musical training shapes the brain. *Frontiers in Psychology, 4,* 713. https://doi.org/10.3389/fpsyg.2013.00713

Baudeat, A. (2017). The emergence of language in the hominin lineage: Perspectives from fossil endocasts. *Frontiers in Human Neuroscience, 11,* 427. https://doi.org/10.3389/fnhum.2017.00427

Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011). Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch. *Brain and Cognition, 77*(1), 1–10. https://doi.org/10.1016/j.bandc.2011.07.00

Boersma, P., & Weenink, D. (2019). *Praat: Doing phonetics by computer* (Version 5.4.01) [Computer software]. http://www.praat.org/

Brandt, A., Gebrian, M., & Slevc, L. R. (2012). Music and early language acquisition. *Frontiers in Psychology, 3,* 327. https://doi.org/10.3389/fpsyg.2012.00327

Brown, S. (2001). The "musilanguag" model of music evolution. In S. Brown, B. Merker, & N. Wallin (Eds.), *The origins of music* (pp. 271–301). MIT Press.

Brown, S., Martinez, M., & Parsons, L. (2006). Music and language side by side in the brain: A PET study of the generation

of melodies and sentences. *European Journal of Neuroscience, 23*(10), 2791–2803. https://doi.org/10.1111/j.1460-9568.2006.04785
Casillas, J. (2015). Production and perception of the /i/-/I/ vowel contrast: The case of L2-dominant early learners of English. *Phonetica, 72*(2–3), 182–205. https://doi.org/10.1159/000431101
Chobert, J., & Besson, M. (2013). Musical expertise and second language learning. *Brain Sciences, 3*(2), 923–940. https://doi.org/10.3390/brainsci3020923
Council of Europe. (2011). *Common European Framework of Reference for Languages: Learning, teaching, assessment*.
Cruttenden, A. (2014). *Gimson's pronunciation of English* (8th ed.). Routledge. https://doi.org/10.4324/9780203784969
Darcy, I., Ewert, D., & Lidster, R. (2012). Bringing pronunciation instruction back into the classroom: An ESL teachers' pronunciation "toolbox". In J. Levis & K. LeVelle (Eds.), *Proceedings of the 3rd Pronunciation in Second Language Learning and Teaching Conference, Sept. 2011* (pp. 93–108). Iowa State University.
Dawson, C., Aalto, D., Šimko, J., Vainio, M., & Tervaniemi, M. (2017). Musical sophistication and the effect of complexity on auditory discrimination in Finnish speakers. *Frontiers in Neuroscience, 11,* 213. https://doi.org/10.3389/fnins.2017.00213
Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals. Evidence-based perspectives for L2 teaching and research*. John Benjamins. https://doi.org/10.1075/lllt.42
Dogil, G. & Reiterer, S. (Eds.). (2009). *Language talent and brain activity*. De Gruyter Mouton. https://doi.org/10.1515/9783110215496
Dolman, M., & Spring, R. (2014). To what extent does musical aptitude influence foreign language pronunciation skills? A multi-factorial analysis of Japanese learners of English. *World Journal of English Language, 4*(4), 1–11. https://doi.org/10.5430/wjel.v4n4p1
Draxler, J. Ch., & Jänsch, K. (2019). *SpeechRecorder — A universal platform independent multi-channel audio recording Software* [Computer program] (Version 4.4.50). Retrieved September 1, 2019, from https://www.bas.uni-muenchen.de/forschung/Bas/software/speechrecorder/
Duckworth, A., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology, 92*(6), 1087–1101. https://doi.org/10.1037/0022-3514.92.6.1087
Fabricius, A. H., Watt, D., & Johnson, D. E. (2009). A comparison of three speaker-intrinsic vowel formant frequency normalization algorithms for sociophonetics. *Language Variation and Change, 21*(3), 413–435. https://doi.org/10.1017/S0954394509990160
Fadiga, L., Craighero, L., & D'Ausillo, A. (2009). Broca's area in language, action, and music. *Annals of the New York Academy of Sciences, 1169,* 448–458. https://doi.org/10.1111/j.1749-6632.2009.04582.x
Fenk-Oczlon, G. (2017). What vowels can tell us about the evolution of music. *Frontiers in Psychology, 8,* 1581. https://doi.org/10.3389/fpsyg.2017.01581
Flege, J. E. (1995). Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issue in cross-language research* (pp. 229–273). York Press. https://doi.org/10.1177/002383099503800102
Flege, J. E. (2007). Language contact in bilingualism: Phonetic system interactions. In J. Cole & J. Hualde (Eds.), *Laboratory phonology* (Vol. 9, pp. 353–380). Mouton de Gruyter.
Fonseca-Mora, M., Toscano-Fuentes, C., & Wermke, K. (2011). Melodies that help: The relation between language aptitude and musical intelligence. *Anglistik International Journal of English Studies, 22*(1), 101–118.
Franklin, M., Moore, K., Yip, C., & Jonides, J. (2008). The effects of musical training on verbal memory. *Psychology of Music, 36*(3), 353–365. https://doi.org/10.1177/0305735607086044
Gaser, C., & Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *The Journal of Neuroscience, 23*(27), 9240–9245. https://doi.org/10.1523/JNEUROSCI.23-27-09240.2003
Gralińska-Brawata, A., & Rybińska, P. (2017). The relationship between the production of word stress and musical abilities in Polish learners of English. *Research in Language, 15*(3), 265–283. https://doi.org/10.1515/rela-2017-0015
Hyde, K. L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A. C., & Schlaug, G. (2009). Musical training shapes structural brain development. *The Journal of Neuroscience, 29*(10), 3019–3025. https://doi.org/10.1523/JNEUROSCI.5118-08.2009
Ikeda, S., Shibata, T., Nakano, N., Okada, R., Tsuyuguchi, N., Ikeda, K., & Kato, A. (2014). Neural decoding of single vowels during covert articulation using electrocorticography. *Frontiers in Human Neuroscience, 8,* 125. https://doi.org/10.3389/fnhum.2014.00125
Isbell, D. (2016). The perception-production link in L2 phonology. *MSU Working Papers in SLS 2016, 7,* 57–67.
Jilka, M., Lewandowski, N., & Rota, G. (2011). Investigating the concept of talent in phonetic performance. In M. Wrembel, M. Kul, & K. Dziubalska-Kołaczyk (Eds.), *Achievements and perspectives in SLA of speech: New Sounds 2010* (pp. 171–180). Peter Lang.
Jusczyk, P. (1999). Narrowing the distance to language: One step at a time. *Journal of Communication Disorders, 32*(4), 207–222. https://doi.org/10.1016/S0021-9924(99)00014-3
Kartushina, N., & Frauenfelder, U. H. (2014). On the effects of L2 perception and of individual differences in L1 production on L2 pronunciation. *Frontiers in Psychology, 5,* 1246. https://doi.org/10.3389/fpsyg.2014.01246
Kendall, T., & Thomas, E. R. (2010). *Vowels: Vowel manipulation, normalization, and plotting in R*. R package. http://cran.r-project.org/web/packages/vowels/index.html
Kolinsky, R., Pascale Lidji, P., Peretz, I., Besson, M., & Morais, J. (2009). Processing interactions between phonology and melody: Vowels sing but consonants speak. *Cognition, 112*(1), 1–20. https://doi.org/10.1016/j.cognition.2009.02.014
Kunert, R., Willems, R. M., Casasanto, D., Patel, A. D., & Hagoort, P. (2015). Music and language syntax interact in Broca's area: An fMRI study. *PLOS ONE, 10*(11). https://doi.org/10.1371/journal.pone.0141069
Lappe, C., Herholz, S., Trainor, L., & Pantev, C. (2008). Cortical plasticity induced by short-term unimodal and multimodal musical training. *The Journal of Neuroscience, 28*(39), 9632–9639. https://doi.org/10.1523/JNEUROSCI.2254-08.2008
Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods, 44*(2), 325–343. https://doi.org/10.3758%2Fs13428-011-0146-0
Lidji, P., Jolicoeur, P., Régine Kolinsky, R., Moreau, P., Connolly, J. F., & Peretz, I. (2010). Early integration of vowel and pitch processing: A mismatch negativity study. *Clinical Neurophysiology, 121*(4), 533–541. International Speech Communication Association. https://doi.org/10.1016/j.clinph.2009.12.018
Llanes-Coromina, J., Prieto, P., & Rohrer, P. (2018). Brief training with rhythmic beat gestures helps L2 pronunciation in a reading aloud task. In K. Klessa, J. Bachan, A. Wagner, M. Karpiński, & D. Śledziński (Eds.), *9th International Conference on Speech Prosody* (pp. 498–502). https://doi.org/10.21437/SpeechProsody.2018-101

Maddieson, I. (2013). Vowel quality inventories. In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online*. Max Planck Institute for Evolutionary Anthropology.
Magne, C., Jordan, D., & Gordon, R. (2016). Speech rhythm sensitivity and musical aptitude: ERPs and individual differences. *Brain and Language, 153–154,* 13–19. https://doi.org/10.1016/j.bandl.2016.01.001
Mandell, J. (2009). *Electronic music and medical education*. Accessed September 1, 2019, from http://jakemandell.com
Mandell, J., Schulze, K., & Schlaug, G. (2007). Congenital amusia: An auditory-motor feedback disorder. *Restorative Neurology and Neuroscience, 25*(3–4, 334), 323.
Masataka, N. (2007). Music, evolution and language. *Developmental Science, 10*(1), 35–39. https://doi.org/10.1111/j.1467-7687.2007.00561.x
McAuliffe, M., Socolof, M., Mihuc, S., Wagner, M., & Sonderegger, M. (2017). Montreal Forced Aligner: Trainable text-speech alignment using Kaldi. In F. Lacerda (Eds.), *Proceedings of the 18th Conference of the International Speech Communication Association* (pp. 498–502). International Speech Communication Association. http://doi.org/10.21437/Interspeech.2017-1386
Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research, 219*(1–2), 36–47. https://doi.org/10.1016/j.heares.2006.05.004
Milovanov, R., Pietilä, P., Tervaniemi, M., & Esquef, P. A. (2010). Foreign language pronunciation skills and musical aptitude: a study of Finnish adults with higher education. *Learning and Individual Differences, 20*(1), 56–60. http://dx.doi.org/10.1016/j.lindif.2009.11.003
Mithen, S. (2005). *The singing neanderthals: The origins of music, language, mind and body*. London Weidenfeld & Nicolson.
Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., & Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex, 19*(3), 712–723. https://doi.org/10.1093/cercor/bhn120
Nikolsky, A. (2015). Evolution of tonal organisation in music mirrors symbolic representation of perceptual reality. Part-I: Prehistoric. *Frontiers in Psychology, 6,* 1405. https://doi.org/10.3389/fpsyg.2015.01405
Ott, C. G., Langer, N., Oechslin, M. S., Meyer, M., & Jäncke, L. (2011). Processing of voiced and unvoiced acoustic stimuli in musicians. *Frontiers in Psychology, 2,* 195. https://doi.org/10.3389/fpsyg.2011.00195
Pastuszek-Lipińska, B. (2008). Musicians outperform nonmusicians in speech imitation. *Lecture Notes in Computer Science, 4969,* 56–73. https://doi.org/10.1007/978-3-540-85035-9_4
Patel, A. D. (2008). *Music, language, and the brain*. Oxford University Press.
Patel, A. D. (2012). Language, music, and the brain: A resource-sharing framework. In P. Rebuschat, M. Rohrmeier, J. Hawkins, & I. Cross (Eds.), *Language and music as cognitive systems* (pp. 204–223). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199553426.003.0022
Picavet, F., Auberge, V., & Rossato, S. (2012). Can a guided rhythmic approach contribute to the oral performance of learners of L2 English? A case study. In M. Busa & A. Stella (Eds.), *Methodological perspective on second language prosody. Papers from ML2P 2012* (pp. 73–77). Cleup.
Reddy, S., & Stanford, J. (2015). A web application for automated dialect analysis. In M. Gerber, C. Havasi, & F. Lacatusu (Eds.), *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations* (pp. 72–75). Association for computational Linguistics. https://doi.org/10.3115/v1/N15-3015
Rojczyk, A., & Porzuczek, A. (2012). Selected aspects in the acquisition of English phonology by Polish learners—Segments and prosody. In D. Gabryś-Barker (Ed.), *Readings in second language acquisition* (pp. 93–120). Wydawnictwo Uniwersytetu Śląskiego.
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. https://doi.org/10.3389/fnhum.2016.00288
Rosenfelder, I., Fruehwald, J., Evanini, K., Seyfarth, S., Gorman, K., Prichard, H., & Yuan, J. (2014). *FAVE (Forced Alignment and Vowel Extraction) Program Suite v1.2.2*. http://doi.org/10.5281/zenodo.22281
Russo, F. A., Vuvan, D. T., & Thompson, W. F. (2019). Vowel content influences relative pitch perception in vocal melodies. *Music Perception: An Interdisciplinary Journal, 37*(1), 57–65. http://doi.org10.1525/mp.2019.37.1.57
Schellenberg, E. G. (2004). Music lessons enhance IQ. *Psychological Science, 15*(8), 511–514. https://doi.org/10.1111/j.0956-7976.2004.00711.x
Schneider, P., Scherg, M., Dosch, H. G., Specht, H. J., Gutschalk, A., & Rupp, A. (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nature Neuroscience, 5*(7), 688–694. https://doi.org/10.1038/nn871
Schwartz, G. (2010). Rhythm and vowel quality in accents of English. *Research in Language, 8,* 135–147. https://doi.org/10.2478/v10015-010-0011-8
Seashore, C. E., Lewis, D., & Saetveit, J. (1960). *Seashore measures of musical talents*. The Psychological Corporation.
Slevc, L., & Miyake, A. (2006). Individual differences in second language proficiency: Does musical ability matter. *Psychological Science, 17,* 675–681. https://doi.org/10.1111%2Fj.1467-9280.2006.01765.x
Sobkowiak, W. (2008). *English phonetics for Poles* (3rd ed.). Wydawnictwo Poznańskie.
Strait, D., Parbery-Clark, A., Hittner, E., & Kraus, N. (2012). Musical training during early childhood enhances the neural encoding of speech in noise. *Brain and Language, 123*(3), 191–201. https://doi.org/10.1016/j.bandl.2012.09.001
Swaminathan, S., & Schellenberg, E. (2019). Musical ability, music training, and language ability in childhood. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. https://doi.org/10.1037/xlm0000798
Tallal, P., & Gaab, N. (2006). Dynamic auditory processing musical experience and language development. *Trends in Neuroscience, 29*(7), 382–390. https://doi.org/10.1016/j.tins.2006.06.003
Tankus, A., Fried, I., & Shoham, S. (2012). Structured neuronal encoding and decoding of human speech features. *Nature Communications, 3,* 1015. https://doi.org/10.1038/ncomms1995
Torppa, R., & Huotilainen, M. (2019). Why and how music can be used to rehabilitate and develop speech and language skills in hearing-impaired children. *Hearing Research, 380,* 108–122. https://doi.org/10.1016/j.heares.2019.06.003
Trehub, S. E., Schellenberg, G., & Kamenetzky, S. B. (1999). Infants' and adults' perception of scale structure. *The Journal of Experimental Psychology: Human Perception and Performance, 25*(4), 965–975. https://doi.org/10.1037//0096-1523.25.4.965
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. https://doi.org/10.1016/j.lindif.2010.02.004
Wang, H., Mok, P., & Meng, H. (2016). Capitalizing on musical rhythm for prosodic training in computer-aided language learning.

Computer Speech and Language, 37, 67–81. https://doi.org/10.1016/j.csl.2015.10.002

Weckwerth, J. (2011). English TRAP vowel in advanced Polish learners: Variation and system typology. In W. S. Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 2110–2113). City University of Hong Kong.

Zhang, A., Feng, H., Wang, S., & Dang, J. (2016). *Relationship between perception and production of English vowels by Chinese English learners*. 2016 10th International Symposium on Chinese Spoken Language Processing (ISCSLP), Tianjin (pp. 1–5). https://doi.org/10.1109/ISCSLP.2016.7918479

Zwolińska, E. A. (2008). Polish fundamentals of music curriculum. *The GIML Audea, 13*(1), 7–8.

Zybert, J., & Stępień, S. (2009). Musical intelligence and foreign language learning. *Research in Language, 7*. https://doi.org/10.2478/v10015-009-0007-4

Appendix H: Research article 2 (Jekiel 2022)

Jekiel, Mateusz. 2022. "L2 rhythm production and musical rhythm perception in advanced learners of English", *Poznan Studies in Contemporary Linguistics 58*, 2: 315-340. https://doi.org/10.1515/psicl-2022-0016

Poznań Studies in Contemporary Linguistics 58(2), 2022, pp. 315–340

doi: 10.1515/psicl-2022-0016

L2 RHYTHM PRODUCTION AND MUSICAL RHYTHM PERCEPTION IN ADVANCED LEARNERS OF ENGLISH

MATEUSZ JEKIEL
Adam Mickiewicz University, Poznań
mjekiel@amu.edu.pl

ABSTRACT

The aim of this research is to investigate the relation between musical aptitude and the acquisition of L2 rhythm by Polish advanced learners of English. A longitudinal study was conducted among 50 Polish students of English reading the “Please Call Stella” passage before and after an intensive two-semester accent training course supplemented by an extensive practical course in English phonetics and phonology. Participants also completed two musical hearing tests (Mandell 2009) and a survey on musical experience. Automated alignment was performed in DARLA (Reddy and Stanford 2015) and reviewed in Praat (Boersma and Weenink 2019). We compared the rhythm metrics calculated in Correlatore (Mairano and Romano 2010) before and after training and juxtaposed them against the pronunciation teachers’ results. We reported a significant difference between the scores for vocalic intervals across all rhythm metrics, indicating that participants’ produced higher vocalic variation after training, more similar to their teachers. However, we observed no significant relationship between the participants’ rhythm metric scores and their musical hearing test scores or musical experience, suggesting that musical aptitude might not play a crucial role in the L2 rhythm production in a formal academic learning environment.

KEYWORDS: L2 rhythm; rhythm metrics; accent training; language and music; musical aptitude.

1. Introduction

Linguists and musicologists agree that the relationship between language and music can be observed on many levels: (1) both are specific to humans and appear in every society throughout history (Nettl 2000); (2) they share the same neural resources responsible for processing linguistic and musical structures (Koelsch et al. 2002); (3) both have tone, melody, and rhythm, as well as measurable duration, frequency, and timbre (Fadiga et al. 2009); (4) and they are both constructed out of syntactic sequences of sounds that can be used to produce an infinite number of possible hierarchically structured signals (Fenk-Oczlon 2009). These similarities led to a number of studies interested in establishing whether musical skills can be related to language skills. Both musical aptitude, i.e. the "potential to achieve in music" (Gordon 1989), as well as musical experience, i.e. music education and performance, have been connected to language skills (Magne et al. 2006). Musical aptitude can positively influence processing of speech, both in terms of perception and production (Chobert and Besson 2013); as well as phonological awareness (Culp 2017) and successful second language learning (Milovanov 2004). Formal music education and musical training can also have a positive effect on the brain mechanisms related to language skills (Wong et al. 2007), although it has also been argued that the majority of such studies do not derive their data from experimental studies (see Patel 2008 for fuller treatment).

Recent research suggests that musical aptitude can also have a positive effect on the acquisition of L2 pronunciation, including both segmental and suprasegmental features. In a study by Milovanov et al. (2008), children with superior musical skills acquired advanced English pronunciation, while Slevc and Miyake (2006) confirmed a link between musical aptitude and second language listening discrimination and speech production among Japanese 11-year-old learners of English. In the case of adult language learners, Milovanov et al. (2010) found a relation between musical aptitude and pronunciation skills in Finnish learners of English. Musical aptitude can also affect direct speech imitation (Pastuszek-Lipińska 2008), which constitutes a learning task, as well as accent faking (Coumel et al. 2019), which relies on retrieving previously stored phonetic representations. The above-mentioned studies suggest that investigating the exact interactions between specific aspects of musical aptitude, i.e. pitch perception, melodic memory, and rhythm perception, and specific aspects of L2 pronunciation, i.e. segmental and suprasegmental phonetics, should be regarded as essential in the process of understanding the precise relationship between these two domains.

While language and music share some commonalities with regard to rhythm, recent research still remains inconclusive. Both speech and music have complex hierarchical rhythmic structures and similarly use pitch and duration to mark group boundaries, albeit in a different manner (Patel 2008). Several studies confirmed the existence of shared neurocognitive resources for rhythm in speech and music (e.g. Magne et al. 2016), suggesting that rhythmic grouping in music can be considered as an offshoot of prosodic grouping in speech. Some researchers even suggested that musical rhythm can be a reflection of the musicians' language rhythm, based on similar differences found between the rhythm metric scores for the composer's native language and their instrumental music, e.g. in the case of English and French speech and early 20th century English and French classical music (Patel and Daniele 2003), although these results were later regarded as questionable due to the complex nature of classical music rhythmic patterns and the effect of sample selection on the rhythm metric scores (London and Jones 2010).

This study is particularly interested in the relationship between L2 rhythm and musical rhythm. According to Jusczyk (1999), rhythmic perception plays an important role in successful language acquisition and is related to phonological awareness (Moritz et al. 2013). Llanes-Coromina et al. (2018) also argue that practicing musical rhythm can improve L2 pronunciation, specifically in terms of perceived comprehensibility and fluency, which are often related to perceived language rhythm. Roncaglia-Denissen et al. (2016) suggest an alternative relationship, claiming that second language learners can develop enhanced musical rhythmic perception, especially if the learners' L1 rhythmic properties significantly differ from their L2. In the study, adult Mandarin, Turkish, and Dutch learners of English were compared with Turkish monolinguals in terms of their musical skills using a melodic and rhythmic aptitude test (Wallentin et al. 2010), similar in assumptions and procedures to the ones used in this study. The authors reported a significant difference in the rhythmic aptitude test scores between monolinguals and bilinguals, as well as between the learners whose L1 rhythm is traditionally considered as different than English, e.g. Mandarin learners of English scored better than their Dutch counterparts, although the latter still scored better than Turkish monolinguals. However, to the best of our knowledge, there is still no study that investigates the relationship between musical aptitude and the acquisition of L2 rhythm using language rhythm metrics. This study will try to enrich the current knowledge on the relation between L2 rhythm and musical rhythm, using quantitative rhythm metrics and musical aptitude measurements.

1.1. Language rhythm and rhythm metrics

Although there is a plethora of studies covering various intricacies of language rhythm, there is still no universally accepted definition of this phenomenon. According to Dellwo (2006), rhythm is the systematic organisation of prominent and less prominent speech units in time, where speech units are defined as syllables or vocalic intervals, while prominence as longer duration, stronger intensity or higher frequency. According to Gut (2012), speech rhythm can be regarded as the temporal organisation of language, a periodic recurrence of events or a structural property of a language, which allows to categorise it to a particular rhythm class or locate it on a rhythmic continuum. As all languages are rhythmically organised, researchers have tried for a considerable period of time to label every language in respect to their rhythmic features with various outcomes.

The first attempt to classify languages according to their rhythm was made by Pike (1945) using the concept of isochrony, i.e. the rhythmic division of time into equal portions, which was later supported by a physiological explanation found in Abercrombie (1967). According to this hypothesis, languages can be categorised either as syllable-timed, i.e. having syllables of equal duration (usually Romance languages, e.g. French, Italian, Spanish); or stress-timed, i.e. having patterns of equal duration between stressed syllables (usually Germanic languages, e.g. English, German, Dutch). While such a dichotomy can be regarded as structurally elegant, later studies have shown that many languages are not easily subject to such classification (Ramus et al. 1999), although some perceptual studies also confirmed that listeners can indeed discriminate between selected stress-timed and syllable-timed languages (Nazzi et al. 1998, Ramus et al. 2003). Therefore, there was a need to find acoustic correlates of language rhythm in the speech signal through experimental studies. Roach (1982) suggested that languages categorised as stress-timed would need to have significant variation between stressed and unstressed syllables and rely on vowel reduction to keep the stressed syllables relatively salient, while syllable-timed languages would need to have fairly equal syllable length.

Following these assumptions, Dauer (1983) proposed that rhythm is a gradient feature of a language, rather than an absolute one, and results from a combination of phonetic, phonological, lexical and syntactic facts, such as syllable structure, word stress and vowel reduction. The model proposed by Dauer is also considered as the precursor for the development of rhythm metrics, which allow to conduct quantitative research into language rhythm and measure the rhythmic differences between languages belonging to the conventional rhythmic categories. These rhythm metrics rely on the timing relations between vocalic and consonantal[1] intervals, assuming that language rhythm is

[1] Some studies use the term "intervocalic" instead of "consonantal". For consistency with rhythm metric names, we use the latter.

primarily determined by vocalic and consonantal variation. For example, stress-timed languages such as English or German allow vowel reduction and complex consonant clusters, which lead to higher variation of content of vocalic and consonantal intervals, while syllable-timed languages usually do not have these features, and consequently, have lower vocalic and consonantal variation.

The first rhythm metrics proposed by Ramus et al. (1999) were used to divide speech into vocalic and consonantal intervals in order to calculate the proportion of these intervals within a sentence and their standard deviation. %V and %C represent the timing proportions of vocalic and consonantal intervals in a single utterance, while ΔV and ΔC (also referred to as Vdev and Cdev respectively) stand for the standard deviations of these intervals. The results of these metrics for several languages corresponded to the traditional rhythm categories proposed by Pike (1945) and Abercrombie (1967). In a subsequent study by Ramus et al. (2003), these metrics have been validated in a perceptual study, where participants listening to modified speech with degraded segments and intonation could discriminate between English and Polish, which had different %V and ΔC scores, but couldn't differentiate between English and Dutch, which had similar rhythm metric scores.

Although %V is regarded as a fairly reliable predictor of rhythm types, ΔV and ΔC results can be easily affected by speech rate variation. To resolve this issue, Dellwo and Wagner (2003) and Dellwo (2006) introduced the rate-normalised metrics VarcoV and VarcoC in order to normalise the results across fluctuating speech rates by means of calculating the standard deviation of vocalic/consonantal interval duration, dividing it by the mean vocalic/consonantal interval duration and multiplied by 100. Dellwo and Wagner (2003) suggested that VarcoC can be more accurate than ΔC in discrimination between stress-timed and syllable-timed languages, while White and Mattys (2007) confirmed that VarcoV can be more precise than ΔV in capturing rhythmic differences between languages. While normalisation can provide more accurate results by neutralising differences in speech rate, Mairano and Romano (2011) also point out that this can potentially neglect relevant phenomena connected to speech tempo, thus it is recommended to use both raw and normalised rhythm metrics in order to obtain a more reliable and complete profile of the investigated material.

Another series of rhythm measurements was proposed by Low et al. (2000) and Grabe and Low (2002). The PVI (Pairwise Variability Indices) are measures of temporal patterning in speech used for comparing languages by means of calculating the degree of durational contrast between successive elements in a sequence. These metrics can be used to compare adjacent vocalic (PVI-V) or consonantal (PVI-C) intervals, providing the mean of the differences between these successive intervals divided by their sum. The PVI measures were designed respectively to reflect alternating stressed syllables and consonant cluster variation, which are considered as the two most important features of stress-timed languages. While a higher vocalic PVI result indicates higher variation in vowel duration and the presence of vowel reduction, a higher consonantal PVI result implies higher variation of consonantal clusters. The nPVI (normalised Pairwise Variability Indices) measures durational contrasts between consecutive elements in a sequence and normalises them for fluctuations in speech rate. Specifically, this rhythm metric calculates the variation in length of neighbouring pairs of vocalic and consonantal intervals and averages these results over the entire speech. Similarly to the rate-normalised VarcoV and VarcoC, nPVI-V and nPVI-C can avoid the distortion of results due to speech rate variation. These metrics were successfully used to study the rhythmic differences between accents of English, confirming that Singaporean English is more syllable-based than stress-based British English (Low et al. 2000). Although the difference between the two accents is relatively small when compared with other languages, it shows that rhythm metrics can be used to detect finer rhythmic characteristics both across languages and accents. Another method for normalising the PVI scores is the CCI (Control/Compensation Index) in Bertinetto and Bertini (2008), which divides the duration of each interval by the number of segments in it. Similarly to the PVI measures, CCI-V is used for vocalic intervals and CCI-C for consonantal intervals respectively.

In order to guarantee reliable and reproducible rhythm metric results, certain measures need to be implemented. First, it is important to rely on a consistent segmentation procedure, as the segmentation choices can affect the results (Mairano and Romano 2011). Automatic segmentation should provide more uniform results and facilitate the preparation of larger speech samples, as previous studies relied on a limited number of speakers due to time-consuming manual segmentation. Second, the use of one type of speech data is crucial, as rhythm metric results can significantly vary between spontaneous and read speech (Dellwo and Wagner 2003, Arvaniti 2012). Moreover, the structure of the reading text alone can influence the results, as Gibbon (2003) confirmed that the same ratio of alternating patterns and monotonic geometric series can yield the same rhythm metric scores. Finally, a common issue in studies on speech rhythm stems from speech rate, as the speed of articulation can affect the results of some rhythm metrics and reveal that rhythmic variability within a particular language can be as significant as between languages (Arvaniti 2012). For example, this can be a result of an unnatural, slowed down manner of reading a text by L2 speakers. Therefore, it is crucial to either control for speech rate or avoid comparing L2 speech to L1, as it might yield distorted results.

1.2. Rhythm in English and Polish

English has been traditionally classified as a prototypical stress-timed language, primarily due to its complex syllable structure, vowel reduction and variable stress patterns. On the other hand, the status of Polish rhythm category has been changing over the years. Rubach and Booij (1985) argued that Polish is an atypical case of a stress-timed language, while Avery and Ehrlich (1992) claimed that it should be regarded as a syllable-timed language due to its lack of vowel reduction. Alternatively, Polish has been categorised as a mixed-type language, occupying an intermediate position and exhibiting features of both stress-timed languages (variable and complex syllable types) and syllable-timed languages (lack of phonological vowel reduction) (Nespor 1990, Grabe and Low 2002). In a recent study by Mairano and Romano (2011), Polish was also considered as a rather isolated case across other European languages, having both low vocalic variability due to the lack of vowel reduction and high consonantal variability caused by complex consonant clusters. Due to these measurable differences between Polish and English rhythm, it should be possible to observe non-native rhythm metric scores in Polish learners of English when they are speaking English. Specifically, we expect that the L2 English rhythm metric scores will be closer to L1 Polish than L1 English due to inconsistent vowel reduction and vowel duration, the two key features that significantly affect rhythm metric scores and are considerably different in Polish and English.

1.3. Teaching second language rhythm

While language rhythm is considered to be one of the first aspects of language acquired by infants (Nazzi et al. 1998), the acquisition of L2 rhythm can be problematic, especially for adult learners, and contributes to the perception of foreign accent (Barry 2007). A number of studies have confirmed the influence of prosodic features on learner's intelligibility and comprehensibility (Avery and Ehrlich 1992, Roach 2002, Field 2005), including the role of L2 speech rhythm. In the case of learners of English as a foreign language (henceforth EFL), non-native rhythm usually results from insufficient durational differences between speech intervals, incorrect word stress and the misapplication of pauses (Adams 1979).

There is still a scarcity of studies on L2 rhythm production, e.g. Guilbault (2002) and Gut (2009) found no effect of pronunciation training or staying abroad on learners' rhythmic properties. A few recent studies (e.g. Lin and Wang 2005, Grenon and White 2008) used different rhythm metrics to compare L2 learners, revealing that their rhythmic scores were located between their L1 and L2 target values. Tortel and Hirst (2010) investigated the rhythmic scores of French learners of English, confirming that rhythm metrics can be successfully used to differentiate between learners and native speakers, as well as learners with varying levels of proficiency, while pointing out the effect of hyperarticulation on the results, a common issue across EFL learners. However, studies on rhythmic differences between learners with different levels of proficiency have been inconclusive. Ordin and Polyanskaya (2015) observed a similar development from syllable-timed patterns to stress-timed speech across L2 learners of English with different L1 rhythm types and various degrees of L2 competence, although German learners of English achieved more native-like rhythm metric scores than corresponding French learners of English, even when the latter group was at an advanced level of proficiency. Jang (2008) found no significant differences between similar groups, while a recent study by White and Mok (2018) reported significant development in L2 rhythm production by observing a change in rhythm metric scores across their subjects, noticing a stronger effect of language experience over length of residence, albeit their sample comprised only five participants.

There is also insufficient research regarding the development of L2 rhythm production among Polish learners of English. Gralińska-Brawata (2014) investigated the effect of pronunciation training on rhythm metric scores in advanced Polish EFL learners. Apparent progress was reported across all participants in at least one rhythm measurement, but only ΔV and VarcoV scores had significant differences between L1 and L2. Although the reliability of rhythm metrics in observing the development of L2 rhythm production still remains inconclusive, the results suggest that language rhythm can be learnable and teachable. Therefore, we expect to notice a difference in the L2 rhythm metric scores before and after training among our participants. Specifically, we expect that the L2 rhythm metric scores will be more similar to native-like results after training due to the acquisition of vowel reduction and greater variation in vowel duration. Moreover, since this study is also interested in the influence of musical aptitude on second language pronunciation, we assume to report more native-like L2 rhythm metric scores before and after training among learners with finer musical aptitude. To summarise, this study will try to answer the following questions:

1. Does accent training affect rhythm metric scores in Polish advanced speakers of English? If yes, which rhythm metrics can indicate this change?
2. Are rhythm metrics capable of differentiating between intermediate and advanced language learners' L2 rhythm production?
3. Does musical aptitude, particularly rhythm perception, affect rhythm metric scores in L2 speech?

2. Method

2.1. Participants

The following study comprised 50 (42 female, 8 male; age 19–21) advanced Polish learners of English (between B2 and C1 according to CEFR framework) with similar results in their secondary school exit exam in English. Before the study, all participants completed LexTALE, i.e. Lexical Test for Advanced Learners of English (Lemhöfer and Broersma 2012), to verify their current level of proficiency in English. The mean result is 74.48% (*SD* = 8.93), which indicates that the group is relatively advanced and homogenous. We observed no L1 accent variation in our participants, i.e. prior to recording English speech, we evaluated impressionistically their Polish spontaneous speech. All participants were using standard Polish (Jassem 2003, Gussmann 2007). Participants had no medically documented speech or hearing impairment and did not have formal accent training prior to the study, i.e. they spoke without a distinctive British or American accent. All participants signed an informed written consent prior to study enrollment and were not remunerated after the study, but received extra course credit for taking part in the study.

2.2. Training

All participants were students of English at university level and received an accent training course in English with a total of 90 hours of class work during the first two semesters as part of their curriculum. The main goal of the course is to teach the students to speak with a native-like General British accent, i.e.

the Standard British English accent spoken in the South of England (students can choose whether to practise General British or General American pronunciation prior to the course start date). The course comprised segmental (i.e. vowels and consonants of English) and suprasegmental phonetics (i.e. intonation, rhythm, syllable stress and sentence stress). The teaching methodology was holistic, i.e. individual features were taught together with prosodic and connected speech processes, rather than in isolation. In the classroom, participants were taught with the use of exercises, reading passages and drama performance. Outside the classroom, participants had to prepare recordings based on various reading texts, dialogues and news reports. During the accent training course, all participants were familiarised with the differences in English vowel length (including such phonetic processes as clipping, i.e. shortening the articulation of vowel before fortis consonants in a stressed syllable) and vowel reduction. These features were practiced in various modes of speech, e.g. via wordlists and reading passages. The participants belonged to four groups and received instruction from four different Polish pronunciation teachers who specialise in teaching English pronunciation and spoke with native-like General British accents. While teaching English pronunciation by non-native speakers can be questionable, the participants should benefit from having instructors with personal experience in learning English pronunciation. Moreover, Polish teachers are aware of typical pronunciation errors made by Polish learners of English and capable of providing sufficient feedback so that students can improve. One of the most common errors made by Polish learners of English is lack of vowel reduction, which is taught during the first academic year and should affect the rhythm metric results (see Sobkowiak 2004 for a comprehensive list). All participants also attended a supplementary practical course in English phonetics and phonology with a total of 45 hours of class work during the first two semesters. The aim of the course is to increase learners' phonological awareness and provide theoretical knowledge for potential future teachers and researchers of English. The course syllabus included articulatory and acoustic phonetics, phonemic and phonetic transcription, intonation and rhythm. Students were assessed via weekly quizzes and four tests.

2.3. Data

A recording of the text "Please Call Stella" was obtained before and after the two-semester accent training course. Both recording sessions were conducted in an anechoic room with an MXL microphone connected to a MacBookPro via a Roland Duo Capture EX audio interface recording in mono 44.1 kHz frequency and 16-bit resolution. The text was displayed on an external monitor in a large white font against a black background. The same procedure was used to record all four pronunciation teachers who conducted the accent training course. The primary reason for analysing L2 English without L1 Polish is related to the fact that native speakers usually have faster speech rate, which in turn can affect the rhythm metric scores, hence the comparison of L2 English and L1 Polish might yield inaccurate results (White and Mattys 2007).

After the first recording session but before the accent training course, all participants completed two musical hearing tests designed by Mandell (2009): the tone deaf test and the rhythm deaf test. The tone deaf test measures pitch perception and melodic memory by playing 36 pairs of synthesised melodies and asking the listener to decide whether each pair is the same or different. The results are expressed in the percentage of correctly identified tokens and use the following scale: below 70% is low, between 70% and 79% is normal, between 80% and 90% is above normal, above 90% is exceptional. The rhythm deaf test measures rhythm perception and rhythmic memory using 25 pairs of synthesised rhythmic patterns and asking the listener to decide whether each pair is the same or different. Similarly to the previous study, the results are expressed in the percentage of correctly identified tokens and use the same scale. Both melodic and rhythmic perception are regarded as good indicators of musical aptitude and both tests are similar in their assumptions and procedures to the ones used in previous studies (Wallentin et al. 2010, Roncaglia-Denissen et al. 2016). Finally, after the musical hearing tests, all participants were asked whether they attended music school or had private music lessons, could play a musical instrument or sing. The answers to these questions were collected via an online survey along with standard demographic data.

After collecting all the data, the recordings were manually edited (i.e. hesitation sounds or repetitions were deleted to avoid their interference in the analysis) and then automatically segmented via FAVE-Extract (Rosenfelder et al. 2014) and Montreal Forced Aligner (McAuliffe et al. 2017) using the DARLA web interface (Reddy and Stanford 2015). Next, all the obtained TextGrid files were verified together with their corresponding audio files and, if necessary, manually corrected using the speech analysis software Praat (Boersma and Weenink 2019). Finally, the rhythm metrics were calculated and plotted on graphs with Correlatore (Mairano and Romano 2010), a programme for rhythmic analysis of annotated Praat TextGrid files. The following rhythm metrics were obtained: Vdev, Cdev, VarcoV, VarcoC, nPVI-V, nPVI-C, CCI-V and CCI-C. All rhythm metrics were computed globally, i.e. including all the values for the consonantal and vocalic intervals in each TextGrid file at once.

3. Results

3.1. Rhythm metrics

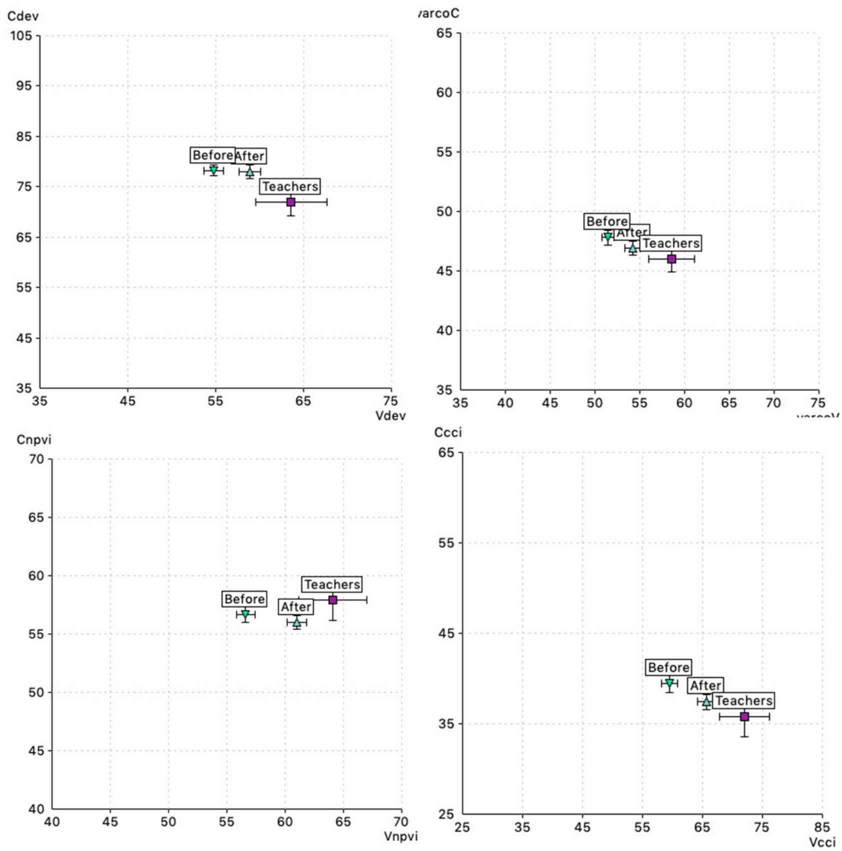


Figure 1. Mean results of participants before and after training with respect to their teachers for individual rhythm metrics: deltas (top left, measuring the standard deviation of V and C intervals), Varco (top right, adding normalisation with respect to speech rate), nPVI (bottom left, measuring the variability of V and C intervals by also considering their succession in time) and CCI (bottom right, measuring the amount of segmental compensation realized by speakers for V and C).

The results presented in Figure 1 show that all rhythm metrics noticed a difference in the participants' rhythmic patterns before and after training. In all cases, the participants' rhythm scores became more similar to their teachers' results after training, taking an intermediate place between the scores before training and the teachers. In terms of language proficiency, this indicates that learners produced more native-like vowel duration contrasts to differentiate between short and long vowels (e.g. short KIT in *bring* and long FLEECE in *please*), as well as produced greater contrasts between stressed and unstressed syllables by using weak forms in their speech (e.g. reducing *to* to /tə/ or *from* to /frəm/). By mastering these characteristic features of English pronunciation, learners' produced more native-like L2 rhythm, since both vowel duration and vowel reduction have a direct impact on English rhythmic patterns.

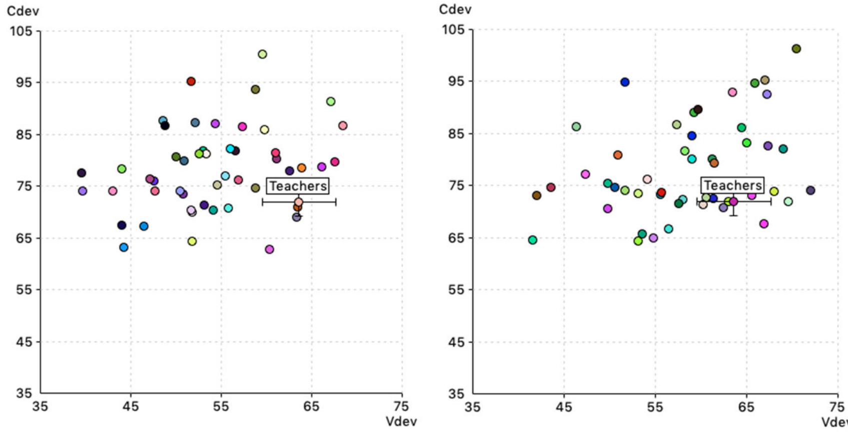


Figure 2. Individual results of participants before (left) and after (right) training with respect to their teachers' mean results for ΔV and ΔC (measuring the standard deviation of V and C intervals).

Figure 2 presents the different ΔV and ΔC scores before and after training. On average, the participants had higher ΔV scores after training ($M = 58.94$, $SE = 1.17$) than before ($M = 54.78$, $SE = 1.15$). Data distribution was normal (before training $W(50) = 0.97$, $p = .23$, after training $W(50) = 0.98$, $p = .79$). A dependent one-tailed *t*-test was performed to compare the rhythm metric scores before and after training. The difference was significant ($t(49) = 3.34$; $p < .01$). However, there was no observable change in the ΔC score before ($M = 78.12$, $SE = 1.15$) and after training ($M = 78.02$, $SE = 1.27$). These results suggest that participants generally improved in vowel reduction and produced more variable vowel duration, while their consonantal variation remained unchanged.

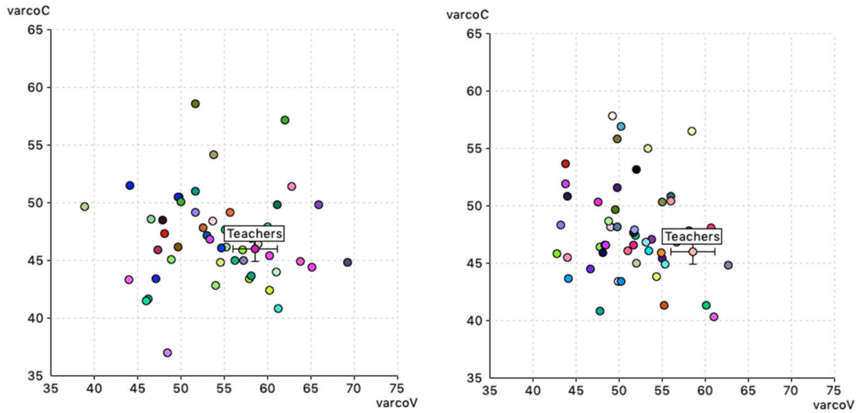


Figure 3. Individual results of participants before (left) and after (right) training with respect to their teachers' mean results for VarcoV and VarcoC (measuring the standard deviation of V and C intervals while adding normalisation with respect to speech rate).

Figure 3 presents the different VarcoV and VarcoC scores before and after training. Participants scored higher in VarcoV after training (*M* = 54.34, *SE* = 0.91) than before (*M* = 51.48, *SE* = 0.70). Data distribution was normal (before training *W*(50) = 0.96, *p* = .16, after training *W*(50) = 0.99, *p* = .95). The difference between the results was significant (*t*(49) = 3.38; *p* < .01). Similarly to ΔC, there was no observable difference for VarcoC before (*M* = 47.83, *SE* = 0.59) and after training (*M* = 46.93, *SE* = 0.56). Overall, we can observe more dispersed results before training than after training, while the results after training are closer to the teachers' results.

The nPVI results presented in Figure 4 show a similar case to the previous rhythm metrics, with more scattered vocalic results before training (*M* = 56.61, *SE* = 0.76), while the same metrics after training are aligned closer to the teachers' scores (*M* = 60.96, *SE* = 0.77). Data distribution was normal (before training *W*(50) = 0.96, *p* = .11, after training *W*(50) = 0.98, *p* = .71). The difference between the results was significant (*t*(49) = 5.21; *p* < .01). Again, there was no observable difference before (*M* = 56.73, *SE* = 0.72) and after (*M* = 56.04, *SE* = 0.59) training for the consonantal nPVI.

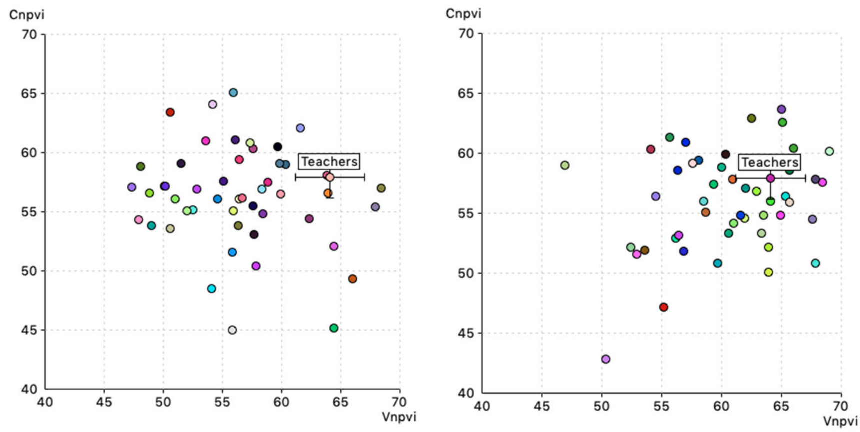


Figure 4. Individual results of participants before (left) and after (right) training with respect to their teachers' mean results for nPVI-V and nPVI-C (measuring the variability of V and C intervals by also considering their succession in time).

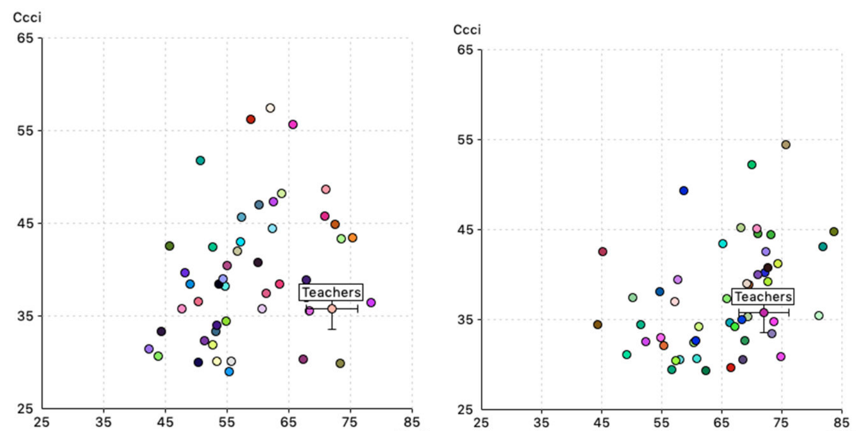


Figure 5. Individual results of participants before (left) and after (right) training with respect to their teachers' mean results for CCI-V and CCI-C (measuring the amount of segmental compensation realised by speakers for V and C).

Finally, the CCI results in Figure 5 show a similar pattern to the previous rhythm metrics, where the vocalic scores after training ($M = 65.55$, $SE = 1.42$)

are closer to the teachers' scores than before (M = 59.50, SE = 1.45). Data distribution was normal (before training $W(50) = 0.96$, $p = .06$, after training $W(50) = 0.97$, $p = .39$). The difference between the results before and after training was significant ($t(49) = 4.13$; $p < .01$). Interestingly, we also reported a significant difference for the consonantal scores before (M = 39.53, SE = 1.02) and after (M = 37.41, SE = 0.85) training ($t(49) = -2.44$; $p < .01$). Since the CCI is designed to measure the amount of compression at a segmental level, languages considered as syllable-timed are expected to have comparable variability of vocalic and consonantal segments, while traditionally stress-timed languages have high variability of vocalic segments and low variability of consonantal segments. Therefore, the result for CCI-C suggests that the participants managed to produce more native-like L2 consonant clusters after training with lower consonantal variation, similar to their teachers. This is also in line with previous studies incorporating the CCI metric (Mairano and Romano 2011).

To summarise, all rhythm metrics allowed us to observe a significant progress in the vocalic scores after training, suggesting that accent training can help students in achieving more native-like vowel duration and vowel reduction, which has a direct effect on perceived speech rhythm in English. At the same time, most rhythm metrics observed no significant change in the consonantal scores after training. This can be related to the fact that the difference in the rhythm metric scores for Polish and English consonantal intervals is much smaller than for vocalic intervals (Mairano and Romano 2011: 1320-1321), leading Polish learners of English to focus on acquiring L2 English vowel duration contrasts and mastering vowel reduction. In conclusion, the results confirm that the above-mentioned rhythm metrics can be successfully used to study a change in L2 rhythm production, even in case of upper-intermediate and advanced learners of English after a year of accent training.

3.2. Musical hearing tests and survey results

Figure 6 presents the tone deaf and rhythm deaf test results, which express the percentage of correctly identified tokens. The mean result for the tone deaf test was 68% (min = 44.4%, max = 86.1%, median 69.4%), while the mean result for the rhythm deaf test was 71.2% (min = 48%, max = 92%, median 72%). Both test results had normal distribution (tone deaf $W(50) = 0.96$, $p = .13$, rhythm deaf $W(50) = 0.95$, $p = .054$). By comparing the results with the provided scale, we can say that in the tone deaf test 27 participants scored below

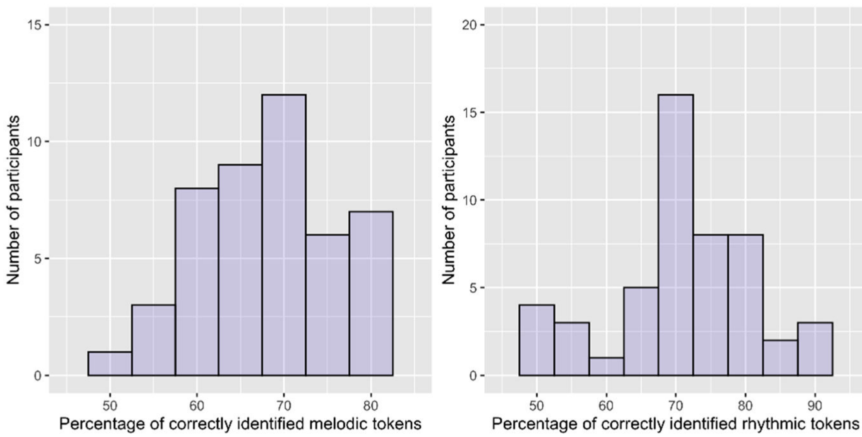


Figure 6. Tone deaf (left) and rhythm deaf (right) test results.

normal (< 70%), 17 normal (70%–79%), and six above normal (80%–90%), while in the rhythm deaf test 20 participants scored below normal (< 70%), 17 normal (70%–79%), 11 above normal (80%–90%), and two exceptional (> 90%).

In the survey, four participants claimed to have completed music school, 12 participants have played a musical instrument, and 12 have sung as soloists, band members or in a choir.

3.3. Musical aptitude and L2 rhythm production

To verify whether musical aptitude is related to L2 rhythm production, we compared the musical hearing test results with the vocalic scores obtained from the rhythm metrics after training. Pearson's Correlation Coefficient reported none of the correlations between the melodic deaf test scores or the rhythm deaf test scores and the individual rhythm metric results after training as significant. All *r* and *p* results are presented in Table 1.

To see if the musical hearing tests correlate with the closeness to the model pronunciation, we also calculated the percentage differences between the participants' rhythm metric scores for vocalic intervals with their teachers' scores and compared them with the participants' musical hearing scores using Pearson's Correlation Coefficient. We observed a weak positive correlation be-

Table 1. Correlation between melodic deaf test (MDT) and rhythm deaf test (RDT) scores and rhythm metric scores before (B) and after (A) training.

		ΔV		**VarcoV**		**nPVI-V**		**CCI-V**	
		B	A	B	A	B	A	B	A
MDT	*r*	−0.08	−0.05	−0.006	0.07	−0.005	0.07	−0.14	−0.10
	p	.58	.73	.96	.62	.97	.62	.33	.48
RDT	*r*	0.15	0.21	0.11	-0.02	0.11	-0.13	0.08	0.13
	p	.28	.14	.44	.89	.44	.36	.54	.36

tween the varcoV scores and the rhythm deaf test results ($r = 0.21$, $df = 49$, $p = .13$), as well as between the nPVI-V scores and the rhythm deaf test results ($r = 0.16$, $df = 49$, $p = .26$). To investigate further, we divided the participants into two groups based on their musical hearing tests (80% being the cut-off point between the below-average and above-average) but found no significant differences in varcoV scores between groups. We also divided the participants into two groups: "musicians", i.e. participants who confirmed in the survey to have completed music school (4) and/or have played a musical instrument (12) and/or have sung as soloists/in a band/in a choir (12), and "non-musicians", i.e. participants without any specific musical experience (33). The result was insignificant for all rhythmic scores after training. These results suggest that musical aptitude or musical experience might not play an important role in the acquisition of L2 rhythm during a one-year accent training course in a formal academic learning environment. However, it is possible that the influence of musical hearing and musical background on the production of L2 rhythm might not be visible due to the method used for assessing L2 rhythm, the relatively short duration of the accent training course, or its formal academic context.

4. Discussion

The acquisition of L2 speech rhythm is a complex issue that can be difficult to measure and is affected by numerous linguistic and extra-linguistic factors.

Our findings suggest that rhythm metrics can be successfully used to observe improvement in L2 rhythm production in a longitudinal study, particularly in regards to increased vocalic variation in L2 English. Interestingly, all rhythm metrics yielded similar results, showing the mean progress of our participants after training and occupying an intermediate place between their rhythm metric scores before training and the mean rhythm metric scores of their pronunciation teachers. Furthermore, the results show that L2 rhythm is teachable and learnable in a formal setting during an accent training course. In this particular case, we observed that Polish adult advanced learners of English changed the degree of their vocalic variation after one year of training, indicating successful acquisition of variable vowel duration and vowel reduction. Finally, we did not find any significant relationship between the musical aptitude results and the rhythm metric scores.

When discussing rhythm metrics, it is vital to remember that these measurements rely on durational properties of speech and do not incorporate other phonological, lexical and syntactic factors affecting perceived speech rhythm (Dauer 1983; Gut 2012). In this regard, language rhythm presented with the use of these metrics is a combination of vocalic and consonantal variation, which are phonetic outcomes of various phonological processes. Therefore, instead of teaching L2 rhythm in a conscious and formal manner with the use of metalanguage, Barry (2007) suggests to concentrate on the underlying phonological processes that affect the perceived speech rhythm in a given language. In the case of EFL pronunciation teaching, the primary focus should be placed on vowel duration and vowel reduction, which have a major effect on English rhythm and are also specifically related to rhythm metric scores. This was also the main method of teaching pronunciation in the accent training course in this study, as pronunciation teachers focused on segmental and suprasegmental phonetics by presenting them in a practical and holistic approach in the classroom, while formal context and metalanguage was discussed in the supplementary course on English phonetics and phonology.

The reason why we did not observe any significant relation between musical hearing tests and L2 rhythm production can be twofold. First, the apparent progress made by our participants might be primarily caused by the accent training course alone, suggesting that all participants can equally benefit from formal instruction when acquiring L2 pronunciation, regardless of their musical aptitude or musical experience. An important factor that should also be addressed is motivation, which can be a valid predictor in successful acquisition of L2 pronunciation among advanced adult learners of English (Smit 2002). Although we tried to control for participants' general proficiency, as well as their performance in and outside the classroom, it is possible that some learners spent more private time on practicing pronunciation. As this was a longitudinal study comparing results before and after one year of studies, individual motivation of our participants could affect the amount of effort they put in during the accent training course, thus, participants with higher motivation could compensate for their lower musical hearing test scores. To verify this in future longitudinal studies, it is crucial to take into account participants' motivation and control for their work outside the classroom. Finally, it is possible that musical hearing can be related to aspects of language rhythm not measured by rhythm metrics, such as f0 or intensity, which also affect perceived speech rhythm (Cumming 2011). Although there is a reported difference between the processing of pitch in language and music (Zatorre and Baum 2012), studies suggest that musical training can improve pitch processing (related to F0) in both domains (Schön et al. 2004). Therefore, it is possible that musical hearing will not directly correlate with rhythm metrics scores, but can still have an effect on successful L2 rhythm acquisition.

The following study has two limitations that should be taken into consideration in future research. First, while the use of rhythm metrics to observe progress across advanced EFL learners proved to be successful, it is crucial to remember that these metrics focus only on the segmental distribution and durational contrasts, omitting other features that can influence language rhythm, such as pitch (Pickering and Wiltshire 2000), sonority (Galves et al. 2002), and loudness (Fuchs 2014). Therefore, it would be valuable to consider adding these features in future studies. Alternatively, it would be beneficial to implement phonological models of rhythm (e.g. Hayes 1995) focusing on prominence patterns and shift towards a more psychological understanding of language rhythm (Arvaniti 2009). Second, while the use of musical hearing tests by Mandell (2009) was a novel approach, it would be also valuable to include tests evaluating music production (Wallentin et al. 2010), as the assessment of actual musical performance (i.e. singing or playing a musical instrument) could provide more authoritative results than participants' self-reported years of musical training. Furthermore, it would be also advantageous to implement tests specifically designed to evaluate participants' rhythmic skills (Bella et al. 2017) and correlate these results with their potential progress in L2 rhythm production. All in all, incorporating these extensions should provide a more complete picture of language rhythm and its potential relation to musical aptitude.

To conclude, this study confirmed the usefulness of rhythm metrics in reporting progress in the acquisition of L2 rhythm in Polish advanced learners of English, complementing earlier studies in this field (e.g. Gralińska-Brawata 2014). At the same time, we did not find any relationship between L2 rhythm metric scores and musical hearing or musical experience. Possibly, musical skills and music education could be related to other factors that constitute speech rhythm and are not expressed by rhythm metrics. Therefore, it is vital to include other factors that constitute L2 speech rhythm, find new ways to quantify language rhythm, as well as include other forms to assess participants' musical hearing and their musical background. We hope that this research will direct future quantitative and longitudinal studies in the field of L2 rhythm production and the role of musical hearing in L2 pronunciation.

5. Acknowledgments

This research was supported by the National Science Centre in Poland, grant number 2014/15/N/HS2/03865.

REFERENCES

Abercrombie, D. 1967. *Elements of general phonetics*. Edinburgh: Edinburgh University Press.

Adams, C. 1979. *English Speech Rhythm and the Foreign Learner*. The Hague: Moulton.

Arvaniti. A. 2009. "Rhythm, timing and the timing of rhythm". *Phonetica* 66(1-2). 46–63.

Arvaniti, A. 2012. "The usefulness of metrics in the quantification of speech rhythm". *Journal of Phonetics*, 40(3). 351–373.

Avery, P. and S. Ehrlich. 1992. *Teaching American English Pronunciation*. Oxford: Oxford University Press.

Barry, W. J. 2007. "Rhythm as an L2 problem. How prosodic is it?". In: Trouvain, J. and U. Gut (eds.) *Non-native Prosody: Bridging the Gap between Research and Teaching*. 97–120.

Bella, D., Farrugia, N., Benoit, C., Begel, V., Verga, L., Harding, E., Kotz, S. 2017. "BAASTA: Battery for the Assessment of Auditory Sensorimotor and Timing Abilities". *Behavior Research Methods* 49. 1128–1145.

Bertinetto, P.M. and C. Bertini. 2008. On modeling the rhythm of natural languages. *Proceedings of Speech Prosody 2008*, Campinas, Brazil. 427–430.

Boersma, P. and D. Weenink. 2019. Praat: doing phonetics by computer [Computer program]. Version 6.1.03, retrieved 1 September 2019 from http://www.praat.org/

Chobert, J. and M. Besson. 2013. "Musical Expertise and Second Language Learning". *Brain Sciences,* 3(2). 923–940.

Cumming R.E. (2011). "The language-specific interdependence of tonal and durational cues in perceived rhythmicality". *Phonetica* 68. 1–25.

Llanes-Coromina, J., P. Prieto and P.L. Rohrer. 2018. "Brief training with rhythmic beat gestures helps L2 pronunciation in a reading aloud task.". *9th International Conference on Speech Prosody 2018*. 498–502.

Coumel, M., M. Christiner and S.M. Reiterer. 2019. "Second Language Accent Faking Ability Depends on Musical Abilities, Not on Working Memory". *Frontiers in psychology* 10(257).

Culp, M.E. 2017. "The Relationship Between Phonological Awareness and Music Aptitude". *Journal of Research in Music Education* 65(3). 328–346.

Dauer R.M. 1983. "Stress-timing and syllable-timing reanalyzed". *Journal of Phonetics* 11. 51–62.

Dellwo, V. and P. Wagner. 2003. "Relations between language rhythm and speech rate". In: Solé, M., D. Recasens and J. Romero (eds.). *Proceedings of the 15th International Congress of Phonetic Sciences*, 3–9 August, Barcelona, Spain. 471–474.

Dellwo, V. 2006. "Rhythm and speech rate: A variation coefficient for deltaC". In: Karnowski, P. and I. Szigeti (eds.). *Language and language processing: Proceedings of the 38th linguistic Colloquium*. Piliscsaba 2003. Frankfurt: Peter Lang. 231–241.

Fabricius, A., D. Watt and D. Johnson. 2009. "A comparison of three speaker-intrinsic vowel formant frequency normalization algorithms for sociophonetics". *Language Variation and Change* 21(3). 413–435.

Fadiga L., L. Craighero and A. D'Ausillo. 2009. "Broca's area in language, action, and music". *Annals of the New York Academy of Sciences* 1169. 44–458.

Fenk-Oczlon, G. and A. Fenk. 2009. "Some parallels between language and music from a cognitive and evolutionary perspective". *Musicae Scientiae* 13(2). 201–226.

Field, J. 2005. "Intelligibility and the Listener: the Role of Lexical Stress". *TESOL Quarterly* 39. 399–423.

Franklin, M., K. Moore, C. Yip and J. Jonides. 2008. "The effects of musical training on verbal memory". *Psychology of Music* 36. 353–365.

Fuchs, R. 2019. "Integrating variability in loudness and duration in a multidimensional model of speech rhythm: Evidence from Indian English and British English". *Speech Prosody* 7. 290–294.

Gabriel, B. 2007. *Learning English Through Songs*. Singapore: Bettyland Publications.

Galves, A., J. Garcia, D. Duarte and C. Galves. 2002. "Sonority as a basis for rhythmic class discrimination". *Proceedings of Speech Prosody 2002*. 323–326.

Gibbon, D. 2003. "Computational modelling of rhythm as alternation, iteration and hierarchy". In: Solé, M., D. Recasens and J. Romero (eds.). *Proceedings of the 15th International Congress of Phonetic Sciences*, 3–9 August, Barcelona, Spain. 2489–2492.

Gordon, E.E. 1989. *Learning sequences in music: Skill, content, and patterns*. Chicago: G. I. A. Publications.

Grabe, E. and E. Low. 2002. "Durational variability in speech and the rhythm class hypothesis". *Laboratory Phonology* 7. 515–546.
Gralińska-Brawata, A. 2014. "Language Experience and Phonetic Training as Factors Influencing Timing Organisation in Polish Learners of English". *Research in Language* 12. 185–198.
Grenon, I. and L. White. 2008. "Acquiring Rhythm: a Comparison of Canadian English and Japanese". In: Chan, H., H. Jacob and E. Kapia (eds.), *BUCLD, Proceedings of the 32nd annual Boston University Conference on Language development*. Somerville, Massachusetts: Cascadilla Press 32. 155–166.
Guilbault, C. 2002. The Acquisition of French Rhythm by Second Language Learners. PhD thesis, University of Alberta.
Gussmann, E. 2007. *The Phonology of Polish*. Oxford: Oxford University Press.
Gut, U. 2009. *Non-native Speech: a Corpus-based Analysis of Phonological and Phonetic Properties of L2 English and German*. Frankfurt: Peter Lang.
Gut, U. 2012. "Rhythm in L2 speech". In: Gibbon, D., D. Hirst and N. Campbell (eds.), *Rhythm, Melody and Harmony in Speech: Studies in Honour of Wiktor Jassem, Special edition of Speech and Language Technology* 14/15. Poznań: Polskie Towarzystwo Fonetyczne. 83–94.
Hayes, B. 1995. *Metrical stress theory: principles and case studies*. Chicago: University of Chicago Press.
Jang, T.Y. 2008. "Speech Rhythm Metrics for Automatic Scoring of English Speech by Korean EFL Learners". *Malsori Speech Sounds. The Korean Society of Phonetic Sciences and Speech Technology* 66. 41–59.
Jassem, W. 2003. "Polish". *Journal of the International Phonetic Association* 33(1), 103–107. https://doi.org/10.1017/S0025100303001191
Jusczyk, P.W. 1999. "Narrowing the distance to language: one step at a time". *Journal of Communication Disorders* 32. 207–222.
Koelsch, S., T.C. Gunter, D.Y. von Cramon, S. Zysset, G. Lohmann and A.D. Friederici. 2002. "Bach speaks: a cortical 'language-network' serves the processing of music". *Neuroimage* 17. 956–966.
Lemhöfer, K. and M. Broersma. 2012. "Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English". *Behavior Research Methods* 44. 325–343.
Lin, H. and Q. Wang. 2005. "Vowel quantity and consonant variance: A comparison between Chinese and English". Paper presented at the Conference Between Stress and Tone, Leiden.
London, J. and K. Jones. 2011. "Rhythmic Refinements to the nPVI Measure: A Reanalysis of Patel & Daniele (2003a)". *Music Perception* 29. 115–120.
Low, E.L., E. Grabe and F. Nolan 2000. "Quantitative characterization of speech rhythm: Syllable-timing in Singapore English". *Language and Speech* 43. 377–401.
Magne, C., D. Schön and M. Besson. 2006. "Musician children detect pitch violations in both music and language better than nonmusician children: behavioral and electrophysiological approaches". *Journal of Cognitive Neuroscience* 18. 199–211.

Magne, C., D.K. Jordan and R.L. Gordon. 2016. “Speech rhythm sensitivity and musical aptitude: ERPs and individual differences”. *Brain and Language* 153/154. 13–19.

Mandell, J. 2009. Electronic Music and Medical Education. (http://jakemandell.com) (date of access: 21 September 2019).

Mairano, P. and A. Romano. 2010. Un confronto tra diverse metriche ritmiche usando Correlatore. In: Schmid, S., Schwarzenbach, M. and D. Studer (eds.). *La dimensione temporale del parlato*, (Proc. of the V National AISV Congress, University of Zurich, Collegiengebaude, 4-6 February 2009), Torriana (RN): EDK, 79–100.

Mairano P. and A. Romano. 2011. “Rhythm metrics for 21 languages”. In: *Proceedings. of the 17th International Congress of Phonetic Sciences*, Hong Kong, China. 1318–1321.

McAuliffe, M., M. Socolof, S. Mihuc, M. Wagner and M. Sonderegger. 2017. “Montreal Forced Aligner: trainable text-speech alignment using Kaldi”. In: *Proceedings of the 18th Conference of the International Speech Communication Association*.

Milovanov, R., M. Tervaniemi and M. Gustafsson. 2004. “The impact of musical aptitude in foreign language acquisition”. In: *Proceedings of the 8th International Conference on Music Perception and Cognition*.

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

Milovanov, R., P. Pietilä, M. Tervaniemi and P.A. Esquef. 2010. “Foreign language pronunciation skills and musical aptitude: a study of Finnish adults with higher education”. *Learning and Individual Differences* 20. 56–60.

Moritz, C., S. Yampolsky, G. Papadelis, J. Thomson and M. Wolf. 2013. “Links between early rhythm skills, musical training, and phonological awareness”. *Reading and Writing: An Interdisciplinary Journal* 26(5). 739–769.

Nazzi T., J. Bertoncini and J. Mehler. 1998. “Language discrimination by newborns: toward an understanding of the role of rhythm”. *Journal of Experimental Psychology: Human Perception and Performance* 24. 756–766.

Nespor, M. 1990. “On the rhythm parameter in phonology”. In: Roca, I. M. (ed.). *Logical Issues in Language Acquisition*, Dordrecht: Foris: 157–175.

Nettl, B. 2000. “An ethnomusicologist contemplates universals in musical sound and musical culture”. In: Wallin, N.L., B. Merker and S. Brown (eds.), *The Origins of Music*. Cambridge. MA: MIT Press. 463–472.

Ordin, M. and L. Polyanskaya. 2015. “Perception of speech rhythm in second language: The case of rhythmically similar L1 and L2”. *Frontiers in Psychology* 6. 1–15.

Pastuszek-Lipińska, B. 2008. "Musicians outperform nonmusicians in speech imitation". *Lecture Notes in Computer Science* 4969. 56–73.

Patel, A. and J.R. Daniele. 2003. “An empirical comparison of rhythm in language and music”. *Cognition* 87(1). 35–45.

Patel, A. 2008. *Music, Language, and the Brain*. New York: Oxford University Press.

Pickering, L. and C. Wiltshire. 2000. “Pitch accent in indian-english teaching discourse”. *World Englishes* 19(2). 173–183.

Pike, K.L. 1945. *The intonation of American English*. Ann Arbor: University of Michigan Press.
Ramus F., M. Nespor and J. Mehler. 1999. “Correlates of linguistic rhythm in the speech signal”. *Cognition* 73. 265–292.
Ramus, F. 2002. “Acoustic correlates of linguistic rhythm: Perspectives”. *Proceedings of Speech Prosody*. 115–120.
Ramus, F., E. Dupoux and J. Mehler. 2003. “The psychological reality of rhythm classes: Perceptual studies”. In: Solé, M., D. Recasens and J. Romero (eds.). *Proceedings of the 15th International Congress of Phonetic Sciences*, 3–9 August, Barcelona, Spain. 337–342.
Reddy, S. and J. Stanford. 2015. “A Web Application for Automated Dialect Analysis”. *Proceedings of NAACL-HLT 2015*.
Roach, P. 1982. “On the distinction between ‘stress-timed’ and ‘syllable-timed’ languages”, in: Crystal, D. (ed.). *Linguistic Controversies: Essays in Linguistic Theory and Practice in Honour of F.R. Palmer*. London. 73–79.
Roach, P. 2002. “Studying rhythm and timing in English speech: Scientific curiosity, or a classroom necessity?”. In: Waniek-Klimczak, E. and P. J. Melia (eds.). *Accents and Speech in Teaching English Phonetics and Phonology. Łódź Studies in Language Vol.5*. Frankfurt am Main: Peter Lang. 199–206.
Roncaglia-Denissen, M.P, D.A. Roor, A. Chen and M. Sadakata. 2016. “The Enhanced Musical Rhythmic Perception in Second Language Learners”. *Frontiers in Human Neuroscience* 10.
Rosenfelder, I., J. Fruehwald, K. Evanini, S. Seyfarth, K. Gorman, H. Prichard and J. Yuan. 2014. FAVE (Forced Alignment and Vowel Extraction) Program Suite v1.2.2 10.5281/zenodo.22281
Rubach, J. and G.E. Booij, 1985. “A grid theory of stress in Polish”. *Lingua* 66. 281–319.
Schön, D., Magne, C., and M. Besson. 2004. “The music of speech: Music training facilitates pitch processing in both music and language”. *Psychophysiology* 41(3). 341–349.
Slevc, L. and A. Miyake. 2006. “Individual differences in second language proficiency: Does musical ability matter?”. *Psychological Science* 17. 675–681.
Smit, U. 2006. “The interaction of motivation and achievement in advanced EFL pronunciation learners”. *IRAL - International Review of Applied Linguistics in Language Teaching* 40(2). 89–116.
Sobkowiak, W. 2004. *English phonetics for Poles*. (3rd edition.) Poznań: Wydawnictwo Poznańskie.
Tortel, A. and D. Hirst. 2010. "Rhythm metrics and the production of English L1/L2". *Speech Prosody 2010*.
Wallentin, M., A.H. Nielsen, M. Friis-Olivarius, C. Vuust and P. Vuust. 2010. “The musical ear test, a new reliable test for measuring musical competence”. *Learning and Individual Differences*. 20. 188–196.
White L. and S.L. Mattys. 2007. “Calibrating rhythm: first language and second language studies”. *Journal of Phonetics* 35. 501–522.
White D. and P. Mok. 2018. “L2 Speech Rhythm Development in New Immigrants”. *9th International Conference on Speech Prosody 2018*. 838–842.

Wong, P., E. Skoe, N. Russo, T. Dees and N. Kraus. 2007. "Musical experience shapes human brainstem encoding of linguistic pitch patterns". *Nature Neuroscience* 10. 420–422.
Zatorre, R. J., and S.R. Baum. 2012). "Musical Melody and Speech Intonation: Singing a Different Tune". *PLoS Biology* 10(7). e1001372.

Address correspondence to:
Mateusz Jekiel
Faculty of English
Adam Mickiewicz University
Collegium Heliodori Święcicki
Grunwaldzka 6
60-780 Poznań
mjekiel@amu.edu.pl

Appendix I: Research article 3 (Jekiel and Malarski 2023)

Jekiel, Mateusz and Kamil Malarski. 2023. "Musical hearing and the acquisition of foreign-language intonation", *Studies in Second Language Learning and Teaching 13*, 1: 151-178. https://doi.org/10.14746/ssllt.23166

Studies in Second Language Learning and Teaching
Department of English Studies, Faculty of Pedagogy and Fine Arts, Adam Mickiewicz University, Kalisz
SSLLT 13 (1). 2023. 151-178
https://doi.org/10.14746/ssllt.23166
http://pressto.amu.edu.pl/index.php/ssllt

Musical hearing and the acquisition of foreign-language intonation

Mateusz Jekiel ✉
Adam Mickiewicz University, Poznań, Poland
https://orcid.org/0000-0002-3906-7793
mjekiel@amu.edu.pl

Kamil Malarski
Adam Mickiewicz University, Poznań, Poland
https://orcid.org/0000-0001-7826-7213
kamil.malarski@amu.edu.pl

Abstract

The present study seeks to determine whether superior musical hearing is correlated with successful production of second language (L2) intonation patterns. Fifty Polish speakers of English at the university level were recorded before and after an extensive two-semester accent training course in English. Participants were asked to read aloud a series of short dialogues containing different intonation patterns, complete two musical hearing tests measuring tone deafness and melody discrimination, and a survey regarding musical experience. We visually analyzed and assessed participants' intonation by comparing their F_0 contours with the model provided by their accent training teachers following ToBI (Tones and Break Indices) guidelines and compared the results with the musical hearing test scores and the survey responses. The results suggest that more accurate pitch perception can be related to more correct production of L2 intonation patterns as participants with superior musical ear produced more native-like speech contours after training, similar to those of their teachers. After dividing participants into four categories based on their musical hearing test scores and musical experience, we also observed that some students with better musical hearing test scores were able to produce more correct L2 intonation patterns. However, students with poor musical hearing test scores and no musical background also improved, suggesting that the acquisition of L2 intonation in a formal classroom setting can be successful regardless of one's musical hearing skills.

Keywords: language and music; intonation; pitch perception; pronunciation learning

1. Introduction

The link between the abilities to produce and perceive speech and music, both unique to humans and universal across cultures, has been extensively studied over the years (see Ott & Jäncke, 2013; Patel, 2008). Superficially, spoken language and instrumental music appear to share a range of characteristic features: both have tone, melody and rhythm, both are organized temporally in syntactically-structured sequences of sounds, and both have a limited number of elements that can be used to form an unlimited number of hierarchically-structured signals (Fenk-Oczlon & Fenk, 2009). The prevalent explanation for these similarities is that the processing of speech and music share common neural networks (Schön et al., 2010). These developments have led to a series of studies attempting to assess possible transfers between these two domains, primarily focused on the relationship between musical skills and language skills. Moreover, various research findings imply that musical hearing and musical experience can improve speech processing, including prosody, segments and syntax (Jentschke & Koelsch, 2009; Schön et al., 2004). Studies on musical skills and second language (L2) learning also suggest that the perception and production of foreign language sounds can be improved by musical ability (Slevc & Miyake, 2006) as well as musical expertise (Chobert & Besson, 2013). Overall, these studies suggest that musical skills and language skills are strongly related and can potentially be employed in the process of second language learning and teaching. What some of the previous studies lack, however, is that learners' musicality was not evaluated empirically, but rather self-reported through questionnaires. Another major limitation has been testing the level of accent proficiency through imitation and shadowing tasks, which too often are not a reliable tool for eliciting the actual second-language accent output. In order to address such limitations, the present study explores the relationship between musical hearing and the acquisition of L2 intonation by Polish advanced learners of English in a longitudinal context. The research aim is to determine whether superior musical hearing skills and musical experience are correlated with more successful production of L2 intonation patterns.

2. Literature review

2.1. Prosody in music and language

While various associations between language and music have been researched through the years, prosody in speech and melody in music are arguably the two most similar aspects, as they both rely on the same acoustic parameters, such as fundamental frequency, amplitude, duration, and spectral characteristics (Schön et al., 2004). One particular part of prosody to be investigated in this study is intonation, that is, the melody of speech resulting from pitch variation used to convey linguistic and pragmatic meaning (Wells, 2014). A widely-discussed process suggesting the existence of shared mechanisms between speech and music is pitch perception. Zatorre and Baum (2012) argue that there are two different pitch processing systems functioning in the brain. The first type is "fine-grained" and it is responsible for the accurate encoding of musical intervals, while the second one is "coarse-grained," and allows for discriminating between different contours in both speech and music. Contour information is also more perceptually salient and can be detected at an early stage by infants, suggesting that it is a more basic and innate process (Chen et al., 2017). Therefore, it is possible to assume that pitch perception in speech and music is related and can have an influence on the language learning process.

Superior perception of contours in music can affect the encoding of contours in speech as a result of common brain functions responsible for processing both of these auditory phenomena (Bidelman & Krishnan, 2009; Wong et al., 2007). Ott and Jäncke (2013) showed that musically trained individuals outperform non-musicians in reaction time for auditory processing of different tonal signals. Behroozmand et al. (2014) found that musicians with absolute pitch (i.e., the ability to label the pitch of a single musical note without the help of a reference sound) develop specialized left-hemisphere mechanisms for pitch processing, unlike musicians with relative pitch (i.e., the ability to discriminate the pitch of a sound after hearing a reference sound) or non-musicians. A series of studies on formal musical practice and tonal language speakers have revealed that musicians learning L2 Mandarin are better at discriminating tone contours in L2 speech than non-musicians (Marie et al., 2010; Wong et al., 2007). This can be related to the fact that first language (L1) Mandarin speakers develop absolute pitch recognition skills in the course of their early language acquisition more often than speakers of non-tone languages (Deutsch, 2002). Zatorre and Baum (2012) argue that not only can musicians be better at encoding tones in speech, but also that tone-language speakers are more accurate in identifying musical tones, suggesting the existence of overlapping cognitive and neural mechanisms.

2.2. Musical hearing and second language intonation

Although a number of articles in popular science discuss the idea of a "musical ear" as an asset in L2 learning, there is no single concept that would encompass an individual's musicality as a whole. Indeed, there are a number of aspects of music that can be related to language learning. First, musical aptitude can be regarded as a talent based on innate motoric and cognitive predispositions and is strongly linked to musical hearing skills, including melodic memory and pitch perception, combined with skills in performing instrumental music or singing (Schön et al., 2004). Secondly, musical experience is related to the perception and production of music through musical training and practice and is independent of musical aptitude (Pastuszek-Lipińska, 2008). In other words, one might have exceptional musical aptitude without performing music, or be a practicing musician in spite of their lack of talent. In this light, musical ability can be regarded as a combination of musical aptitude and musical experience and is expressed in how well an individual can perform music due to their aptitude and/or experience. Musical expertise, in turn, is a broad concept that can be used to characterize an expert in music performance and/or music theory. Both concepts may prove to be an asset for learners acquiring L2 English intonation; therefore, both will be taken into account in the present study.

There are a number of studies confirming the relationship between the processing of pitch in language and music, and its potential impact on listening skills in language learning. Schön et al. (2004) showed that adult professional musicians are more accurate in processing small changes in F_0 in both music and language than non-musicians, with shorter onset latency of the brain waves associated with F_0 manipulations. Based on these results, they argued that extensive musical training can affect the perception of pitch contour in both domains. Follow-up studies by Magne et al. (2006), and Moreno and Besson (2006) corroborated these results, revealing that young musicians outperform non-musicians in detecting pitch violations in speech. For adult listeners, Dankovicová et al. (2007) also found a relationship between musical hearing skills and intonation discrimination among university students of English. Finally, a related study by Patel et al. (2005) confirmed that tone-deaf listeners (i.e., not being able to label the difference between two tones in terms of pitch) have difficulties in discriminating intonation contours in speech. The above-mentioned studies prove that finer musical pitch perception can relate to more accurate identification of intonation patterns in speech.

If pitch perception in music is indeed related to the perception of speech in one's L1, a similar relationship should be found for the perception of L2 speech. Wong et al. (2007) set out to determine this link and observed that musical experience can shape human brainstem encoding of pitch patterns in speech. This was confirmed in a subsequent study by Wong and Perrachione (2007), revealing the relationship between former musical experience and adult learners' identification of non-native pitch patterns. A similar study by Alexander et al. (2005) investigated the discrimination of lexical tones in Mandarin Chinese by adult English-speaking musicians and non-musicians, providing more evidence for greater accuracy in L2 pitch perception and musical experience. On the other hand, Kurt et al. (2014) investigated the effects of explicit instruction and musical experience on the perception of L2 English intonation patterns among international students of English with different L1s (Mandarin, Japanese, Spanish, and Arabic). While the effect of explicit training was found, there was no apparent effect of musical familiarity on the correct identification of L2 intonation patterns. Finally, Intartaglia et al. (2017) compared the listening skills of adult native speakers of English with non-native musicians and non-musicians by recording their subcortical electrophysiological responses. The results of native speakers and non-native musicians were comparable, while non-native non-musicians scored lower, suggesting that musical experience can lead to enhanced neural encoding of acoustic information and compensate for the lack of language experience. Although these studies show that musicians can have an advantage when performing analytic listening tasks in L2, they do not explain whether musical experience can help in the acquisition of L2 pronunciation.

While most former studies have focused on the relationship between pitch perception in music and L2 listening skills, there is still a scarcity of research examining the function of pitch perception in L2 speech production. Slevc and Miyake (2006) found that musical hearing skills are correlated with L2 production skills among Japanese adult learners of English. In Milovanov et al. (2010), Finnish adult students of English with higher scores in *Seashore Measures of Musical Talents* (Seashore et al., 1960) produced fewer errors in a speech shadowing task pronouncing challenging English phonemes (e.g., /ʒ/, /ɜː/, /ð/). Pastuszek-Lipińska (2008) reported that Polish learners of English with formal music education produced fewer errors than non-musicians in a speech shadowing task, although both groups performed at a similar level in terms of intonation. In a related study by Zybert and Stępień (2009), Polish adult learners of English who scored better in Edwin Gordon's *Intermediate Measure of Music Audiation* test also received higher scores from a native speaker in a speech shadowing task focused on intonation, word stress, and overall pronunciation.

One limitation of the above studies stems from the use of speech shadowing tasks, in which participants listen to and repeat isolated words or phrases and, consequently, are restricted to speech imitation and may not necessarily represent learners' actual pronunciation skills (Dufour & Nguyen, 2013; Mitterer

& Müsseler, 2013). Another limitation is related to pronunciation assessment. Instead of using speech analysis software to conduct a more objective assessment, most empirical investigations rely on impressionistic judgements of speech. Finally, as most previous studies have compared pronunciation skills of professional musicians to non-musicians, language learners without formal music education but with good musical hearing are potentially overlooked in such studies (Zarate et al., 2012).

2.3. Intonation in the English as a foreign language (EFL) classroom

Mastering discourse intonation (i.e., the segmenting and topic-structuring function of pitch) in the EFL classroom is one of the most difficult linguistic skills to teach and learn effectively (Roach, 2000) and EFL learners have been frequently reported to produce errors in the realization of various intonation patterns (Willems, 1982). However, it is now agreed that intonation is both teachable and learnable, and that it plays an important role in communication, especially in international settings (Aronsson, 2014).

It is well known (e.g., Pijper, 1983; Willems, 1982) that General British (GB) uses considerably large pitch movements (octave up or down), and most GB intonation patterns start at the mid-level (Nooteboom, 1997), rather than at the bottom level, posing considerable difficulties for non-native speakers whose L1 does not incorporate pitch movements to such an extent. Moreover, learners' intonation is frequently influenced by prosodic patterns in their L1 (Mennen, 2004). Studies have shown that L2 learners have problems with selecting appropriate intonation contours (He et al., 2012), often relying on their native tones instead (Gut, 2009). Grabe and Karpinski (2003) was the first study to provide a prosodically annotated and phonetically descriptive corpus of Polish and English speech data. The analysis confirmed the existence of language-specific properties in intonation as English and Polish speakers produced different nuclear accent types and distributed them differently.

Despite many languages sharing a lot of commonalities in their use of intonation, or prosodic universals, the target classroom L2 pronunciation competences, especially at the proficient level, go beyond these universals. Those language-specific modes of intonation are usually perceived as difficult to teach and learn and, consequently, are often avoided by teachers in the EFL classroom (Demirezen, 2009). While researchers agree that prosody should be formally taught in the EFL classroom (Chapman, 2007), teachers find many aspects of intonation difficult, due to a lack of appropriate materials (Derwing, 2008). Nevertheless, practicing intonation and other suprasegmental features can have a significant effect on spoken proficiency and comprehensibility (Kang, 2013). In our study, we investigate the acquisition of English L2 intonation in a longitudinal context, with a specific focus on how musical hearing and musical experience may influence this process.

3. The present study

While the majority of similar studies have focused on imitation tasks in testing pronunciation and treated self-reports as a measure of the level of musical abilities to determine the link between pitch perception in music and speech, the primary goal of this study is to investigate the relationship between music perception and the production of intonation patterns by Polish advanced learners of English before and after an accent training course. In our study, we try to assess whether musical hearing, as measured by three different tests, translates into better production of L2 intonation patterns after training. Indirectly, we also attempt to measure the extent to which intonation is learnable and teachable by comparing the recorded patterns before and after accent training. Our research questions are as follows:

1) Are participants able to produce more correct intonation patterns after training?
2) Do participants who scored better on the musical hearing tests also produce more correct intonation patterns after training?
3) Do participants with musical experience produce more correct intonation patterns after training, regardless of their musical hearing test results?

4. Method

4.1. Participants

Our participants were 50 Polish university students (42 females, 8 males)[1] of English between the ages of 19 and 21 (*M* = 20.14, *SD* = .40) who spoke with standard Polish intonation in their L1. They were learning English at the advanced level of proficiency, between B2 and C1 within the Common European Framework of Reference (Council of Europe, 2001), and had overall good results in their secondary school exit exams in English (*M* = 87.75, *SD* = 6.56). In order to confirm their language proficiency, we conducted the Lexical Test for Advanced Learners of English (LexTALE) by Lemhöfer and Broersma (2012), which aims to assess general L2 English proficiency. The test results also confirmed their EFL proficiency between B2 and C1 (*M* = 74.48%, *SD* = 8.93). None of the

[1] Language majors in Poland are commonly more popular among female than male applicants.

participants had reported having formal accent training in their English classes at previous stages of their education, nor did any of them have medically documented speech or hearing impairments.

4.2. English accent training course and English phonetics and phonology course

As part of their curriculum, all participants took an obligatory English accent training course, which comprised segmental phonetics (i.e., vowels and consonants of English) and suprasegmental phonetics (i.e., intonation, rhythm, syllable stress and sentence stress). All students attended the course twice a week during the first two semesters (90 hours of class work). The primary objective of the accent training course is to teach the students to speak English with an accent that is as native-like as possible. All participants aimed for a General British pronunciation, that is, the accent spoken in the South of England and the English pronunciation model which is most commonly used in the Polish primary and secondary education system (Weckwerth et al., 2006).[2] The course focused on the pronunciation of GB vowels and consonants, word stress, weak forms, connected speech processes, and intonation.

For the purpose of the study, the participants were divided into four different groups and were taught by four different accent training teachers; that is, academic instructors specializing in teaching English pronunciation to Polish learners of English. All teachers were female Polish speakers of English with near-native GB accent and over 20 years of experience in accent training and L2 research. The primary reason for relying on Polish instructors instead of native speakers of English is that the former have the necessary first-hand experience and insight allowing them to identify the differences between Polish and English pronunciation, which in turn can be used to help the learners avoid potential errors.

Phonetic instruction during both academic semesters was holistic; the focal areas were taught not in isolation, but within the framework of connected speech phenomena of coarticulation, connected speech processes, stress and intonation. Teaching methods included in-class drills and repetitions. Student assessment was performed via weekly in-class drama performance or news-reading and monthly recordings based on authentic materials. During the accent training course, all participants were familiarized with English intonation patterns and practiced their usage in different contexts, usually through dialogue reading and drama performance.

All participants also attended a two-semester practical course in English phonetics and phonology, which supplemented the accent training course by raising

[2] Students are offered a choice between the General American and the General British accent training course.

phonological awareness and making students aware of how English speech sounds are produced and transcribed. All students attended the course once a week during their first two semesters (45 hours of class work). Course topics included articulatory and acoustic phonetics, phonetic transcription, English phonemes and allophones, word stress, connected speech processes, and intonation. Assessment relied on regular weekly quizzes, as well as two mid-semester and two end-semester tests.

4.3. Data collection

4.3.1. Speech production

All participants took part in two recording sessions. The first one took place in the first week of their studies, before they received formal instruction in English phonetics and phonology and accent training, both of which were included in their course of study. The second recording session took place at the end of the second semester, allowing insight into participants' progress. Both recording sessions consisted of a spontaneous conversation in English, followed by reading aloud a set of four short dialogues. The dialogues were adapted from Wells (2014) and were meant to elicit as many different intonation contours as possible (see Appendix for more details). The English part of the interview was preceded by a short spontaneous conversation in Polish to verify any possible speech impediments or dialectal variation in their L1, as Polish intonation can differ in certain regions and could influence the results of our study (Gussmann, 2007). Table 1 summarizes the sentences from the adapted dialogues included in the analyses. It also presents the target intonation contours associated with the sentences, written down in the ToBI convention for transcribing intonation; then it provides the target intonation contour in the nucleus; and in the last column, it includes the function of that intonational phrase (statement, command, tag question etc.).

The intonation patterns were verified on the basis of recordings of the four accent training teachers who had taught the participants, as well as a recording of a native speaker representing the target accent. All teachers were consistent in their production. If two out of four teachers produced a different intonation pattern, we considered it as an acceptable alternative answer (hence two options for *yes/no* questions). A total of 1600 tokens were collected (800 before training and 800 after training). Dialogues were displayed in large black font against white background on a computer screen, one dialogue at a time. A short instruction explaining the task preceded the actual dialogues. The participants were asked to read each dialogue silently and then read each one aloud, trying to sound as natural as possible for the given context. The recordings were obtained in a sound-treated booth using a studio microphone and SpeechRecorder software (Draxler & Jänsch, 2019). The tokens were recorded in mono 44.1 kHz and 16-bit resolution.

Table 1 Sentences used in data analysis

Sentence	ToBI	Nucleus	Function
You will have to CHECK that. I'm going for a jog in the PARK.	H*L-L%	high fall	statement
Let me HAVE some. Hands off my DRINK!	H*L-L%	high fall	command
You're broke again, AREN'T you? So the match is on Sunday, ISn't it?	H*L-L%	high fall	tag question (certainty)
What are you DRINKing? How did you like the football match YESterday?	H*L-L%	high fall	wh-question
Shall I pay for the CLEANing? Did you finish the ESSay?	L*H-H% H*H-H%	low rise high rise	yes/no question
ACtually, I don't really like football. ACtually, let's talk about your homework.	H*L-H%	fall-rise	attitude word
I only want to TASTE it... I've finished the introDUCtion...	H*L-H%	fall-rise	non-finality
NEAT! You'd BETTer!	L+H*L-L%	rise-fall	strong approval

4.3.2. Musical hearing and musical experience

At the end of the first recording session, participants were asked to complete two online musical hearing tests (Mandell, 2009) focusing on tone deafness and melody discrimination. The tests measure important indicators of musical hearing skills and rely on similar rules to other musical hearing tests (see Wallentin et al., 2010). Both tests were conducted in a separate room, using a laptop connected to a pair of closed-back headphones AKG K240 MkII with audio frequency bandwidth 15-25000 Hz. Each participant completed the tests in isolation.

The first test was the *Adaptive Pitch Test*, measuring tone deafness and pitch perception, in which participants listened to a series of two tones (300 ms each, with a 100 ms silence between the first and the second tone) and were asked to determine whether the second tone in each pair was higher or lower than the first one by pressing the UP or DOWN arrow on the keyboard, respectively. Participants could use the spacebar to replay the tones. The next pair was played immediately after providing the answer for the previous pair. The test duration was circa one minute. The test was adaptive, so the number of played tokens varied and relied on the correct answers. The pitch difference between the first two sounds was 96 Hz. After providing three correct responses in a row, the pitch difference was halved from the previous trial to 48 Hz, progressing to the next, more difficult level. After providing an incorrect response, the pitch difference would regress to the previous, easier level. The tones in the test oscillated within the range of 50-500 Hz. The test results, expressed in Hertz (Hz), indicate the accuracy of differentiating between two tones. The data from 11.761 subjects available on the test's website show the normal distribution of data, with the mean result of 3.98 Hz. Table 2 provides the interpretation of the test scores.

Table 2 Interpretation of the *Adaptive Pitch Test* results

Result	Performance
< .75 Hz	Exceptional
< 1.5 Hz	Very good
< 6 Hz	Normal
< 12 Hz	Low-normal
> 16 Hz	Below normal

The second test was the *Tonedeaf Test*, measuring melody discrimination and melodic memory. Its utility in investigating pitch discrimination and musical ability has been verified by Palomar-García et al. (2020) and Ning (2020). Each participant listened to 36 pairs of short (2-8 seconds each, with a 2-second interstimulus interval) instrumental melodies representing various musical styles and had to decide whether the melodies were the same or different by clicking on the corresponding button on the screen. No repetition was possible in this test. Each pair of melodies used different sonorities (e.g., piano, keyboard, string instruments, wind instruments) in order to reduce potential bias due to specific instrument training. Each pair also varied from one another in terms of natural or synthesized sounds, duration, intensity, timbre, and number of tones. 18 pairs were identical, while the other 18 pairs differed in the pitch of one note, which occurred in one of the last ten tones of the melody and was modified by up to 11 semitones. Out of the 18 pairs that were different, half had the different note within the scale of the melody, while the other half had the different note outside of the scale of the melody. The test was designed to assess melodic memory and locate neuroanatomical correlates of tone-deafness (congenital amusia). The test takes five minutes to complete and the results are expressed in percentages, indicating the percent of correctly identified pairs. The data from 61.036 subjects available on the test's website follow a normal distribution, with the mean result of 73.8% (SD = 9.99). Table 3 provides the interpretation of the test scores.

Table 3 Interpretation of the *Tonedeaf Test* results

Result	Performance
> 90%	Exceptional
> 80%	Very good
> 70%	Normal
> 60%	Low-normal
< 55%	Below normal

After the musical hearing tests, we also asked our participants to complete a survey (in Polish) to assess their musical experience. The questionnaire comprised the following questions: Did you go to music school? If yes, when and for how long? Can you play a musical instrument? If yes, what kind of instrument(s) can you play and how long have you been playing? Can you sing? If yes, how often do you sing?

4.4. Data analysis

Intonational phrases (IPs) summarized in Table 1 were extracted from the recordings and analyzed by both authors who are trained phoneticians and active accent training teachers. IPs were transcribed, labeled and analyzed acoustically in Praat (Boersma & Weenink, 2022). We labeled the pitch accents and boundary tones using the ToBI convention (Beckman & Elam, 1997; Brugos et al., 2006); pitch measurements were inspected in Praat using the program's algorithm for fundamental frequency (F_0) tracking with pitch floor set to 75 Hz and pitch ceiling set at 600 Hz. Manual corrections were performed for signal failures, such as octave jumps or pitch halvings. To determine the correct intonation contour, we observed the difference in F_0 between the pitch accent and the boundary tone. Following the established model answers by the teachers and the native speaker, we marked H*L-L% (high fall) as correct patterns for statements, commands, tag questions, and wh-questions; L*H-H% (low rise) or H*H-H% (high rise) for yes/no questions; H*L-H% (fall rise) for attitude words and expressions of non-finality; L+H*L-L% (rise-fall) for strong approval. Each participant could score a point for each correct intonation pattern (two points per function), up to a total of 16 points from one recording session. Figure 1 presents a model IP produced by the teacher, an incorrect IP produced by one participant before training, and a correct IP produced by the same participant after training.

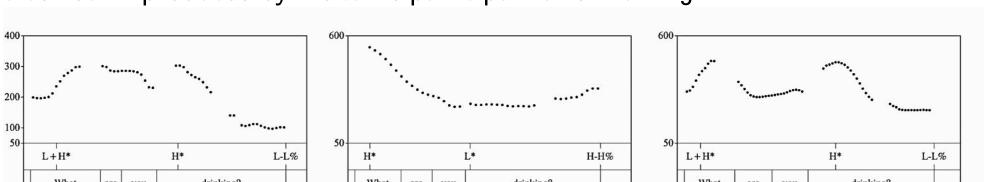


Figure 1 Example wh-question produced by the three speakers: model pattern produced by the teacher (left); incorrect pattern produced by one student before training (mid); correct pattern produced by the same student after training (right)

A one-tailed *t*-test for two dependent means was conducted to investigate participants' progress in acquiring intonation scores before and after training, followed by a linear multiple regression to predict the percentage of correct intonation patterns produced by the participants before and after training, with gender, LexTALE result, *Adaptive Pitch Test* score, *Tonedeaf Test* score, and musical experience as independent variables. Finally, a two-way ANOVA was conducted to examine the effect of musical hearing test scores and musical experience on the production of correct intonation scores before and after training.

5. Results

This section begins with the presentation of intonation scores before and after training, followed by musical hearing test results and participants' musical experience survey responses. Next, we present the effects of phonetic training, musical hearing, and musical experience on intonation scores.

5.1. Intonation scores before and after training

Figure 2 shows an observable improvement in intonation scores for 50 participants after the two-semester accent training course. A one-tailed t-test for two dependent means showed a statistically significant difference, $t(49) = 10.02$, $p < .001$ with 95% CI [-Inf, .15]. The mean result was 49.0% before training ($SD = .13$) and 66.8% after training ($SD = .15$).

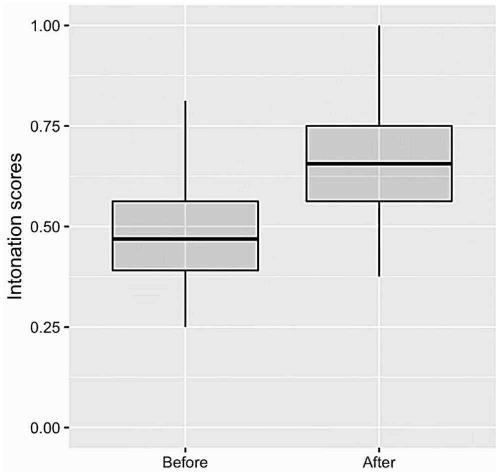


Figure 2 Intonation scores before and after training

Figure 3 shows the overall scores for H*L-L% (high fall) in commands, statements, tag questions, and wh-questions. While statements and commands were both relatively easy to produce for the participants both before (94% and 87%, respectively) and after the training (99% and 95%, respectively), the gains were more pronounced across tag questions (41% before and 66% after) and wh-questions (55% before and 76% after). The most frequent incorrect pattern for these two functions was L*H-H% (low rise) or H*H-H% (high rise).

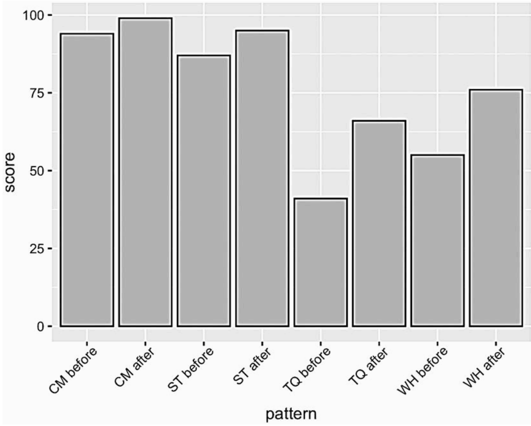


Figure 3 Intonation scores before and after training for H*L-L% (high fall) in commands (CM), statements (ST), tag questions (TQ), and wh-questions (WH)

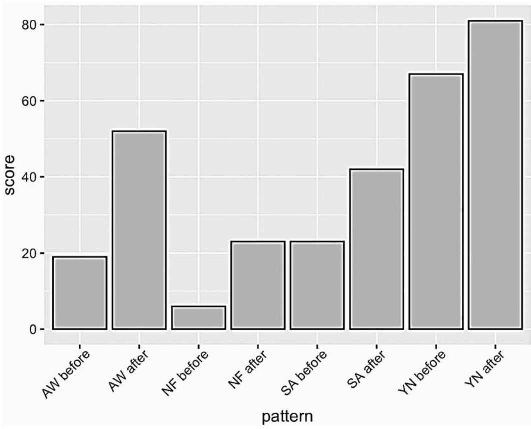


Figure 4 Intonation scores before and after training for the H*L-H% (fall-rise) in attitude words (AW) and non-finality (NF), L+H*L-L% (rise-fall) in strong approval (SA), and L*H-H% (low rise) or H*H-H% (high rise) in yes/no questions (YN)

Figure 4 shows the overall scores for H*L-H% (fall rise) in attitude words and expressions of non-finality; L+H*L-L% (rise-fall) for strong approval; and L*H-H% (low rise) or H*H-H% (high rise) in yes/no questions. The most difficult pattern for participants before training was H*L-H% in non-finality and attitude words (6% and 19%, respectively), although they managed to make a noticeable progress across both functions (23% and 52%, respectively). The second most difficult pattern was L+H*L-L% for strong approval (23% before and 42% after). Out of the following patterns, the use of L*H-H% or H*H-H% for yes/no questions was the least difficult for students, even less difficult than the use of H*L-L% for tag questions or wh-questions (67% before and 81% after). The most frequent incorrect pattern for these functions was H*L-L% (high fall) or L*L-L% (low fall).

5.2. Musical hearing tests

Figure 5 shows the *Adaptive Pitch Test* results, which reveal how precisely participants could differentiate two tones in Hertz values. The mean result was 16.05 Hz (*SE* = 2.24). Participants with musical experience had an average score of 11.08 Hz, while participants without any musical experience had an average score of 18.84 Hz. The two highest results were scored by participants with formal music education and 12 years of musical experience (1 Hz), while the two weakest results were scored by participants without any musical experience (60 Hz).

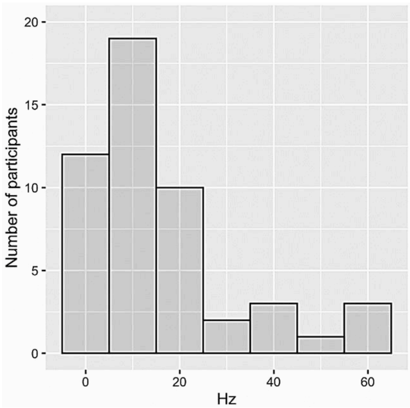


Figure 5 *Adaptive Pitch Test* results

Figure 6 displays the *Tonedeaf Test* results, which express the percentage of correctly identified melodic tokens. The mean result was 68% (*SE* = 1.35). The mean score is similar to the mean score found on the test's website (73.9%) and is comparable to the mean scores found in Ning (2020) for beginner and advanced L2 speakers of Mandarin (63.88% and 74.21%, respectively). Participants with musical experience had an average score of 72%, while participants without any musical experience had an average score of 65.7%. The highest result was scored by a participant with formal music education and 12 years of musical experience (83%), while the weakest result was scored by a participant without any musical experience (44%).

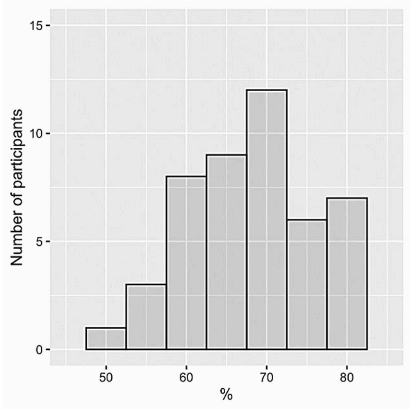


Figure 6 *Tonedeaf Test* results

5.3. Musical experience

Table 4 summarizes the musical experience of our participants that we included as a fixed effect in the linear multiple regression analysis. According to the musical experience survey completed by the participants, 18 out of 50 had some musical experience. Four participants had graduated from music school (first degree); two of them had played a musical instrument for six years, the other two for twelve years. Three participants had played a musical instrument for seven to nine years without any formal music education. Six participants had practiced singing for ten to twelve years without any formal music training. Finally, three participants had played a musical instrument and had practiced singing for four to six years, and two participants had done the same for ten to twelve years. Participants who admitted that they had played a musical instrument or had sung only for a brief episode in their former years (i.e., less than a year) were treated in the analysis as participants with no musical experience.

Table 4 Number of participants with formal music education and musical experience

Years	Music education	Playing a musical instrument (no music education)	Singing (no music education)	Playing a musical instrument and singing (no music education)
10-12	2	-	6	2
7-9	–	3	-	
4-6	2	-	-	3

5.4. Fixed effects on intonation scores before and after training

A linear multiple regression was performed to predict the percentage of correct intonation patterns produced by the participants before the training, based on the following independent variables: gender, LexTALE result, *Adaptive Pitch Test* score, *Tonedeaf Test* score, and musical experience. The regression equation was significant ($F(5,44) = 2.765$, $p < .029$) with an R^2 of .15). The results for individual independent variables are summarized in Table 5 and show that the only significant estimate was found across LexTALE results ($p = .035$), suggesting that a general L2 proficiency might be an indicator in the production of correct intonation patterns. We found no significant results for musical hearing or musical experience at this stage. No multicollinearity between the independent variables in the VIF-scores was uncovered.

Table 5 Summary of fixed effects on intonation scores before training

	Estimate	SE	*z*	*p*	*VIF*
Intercept	.317	.191	-1.656	.105	–
Gender (M)	.098	.051	1.923	.061	1.218
LexTALE	.004	.002	2.172	.035	1.018
Adaptive Pitch Test	*-.002*	.001	-1.317	.195	1.318
Tonedeaf Test	-.002	.002	-.878	.385	1.234
Musical Experience	-.003	.000	-.805	.425	1.337

Note. $N = 50$, $F(5,44) = 2.765$, $p < .029$, $R^2 = .15$, *SE* = Standard Error, *VIF* = Variance Inflation Factor

Table 6 Summary of fixed effects on intonation scores after training

	Estimate	SE	*z*	*p*	*VIF*
Intercept	.655	.235	2.793	.008	–
Gender (M)	.099	.062	1.587	.119	1.219
LexTALE	.002	.002	1.149	.256	1.018
Adaptive Pitch Test	-.003	.001	-1.827	.074	1.318
Tonedeaf Test	-.002	.002	-.999	.323	1.234
Musical Experience	.000	.005	.166	.869	1.337

Note. $N = 50$, $F(5,44) = 1.794$, $p = .134$, $R^2 = .07$, *SE* = Standard Error, *VIF* = Variance Inflation Factor

Another linear multiple regression statistics was run to predict the increase in correct intonation patterns produced by the participants after the training, based on the same independent variables as in the previous analysis. The results for individual independent variables are summarized in Table 6 and show *Adaptive Pitch Test* results ($p = .074$) as significant predictors of intonation score after training. It is important to note that in the case of the *Adaptive Pitch* Test scores, the estimate is negative as higher scores on the test indicate weaker discrimination of pitch in music. Therefore, participants who could discriminate between two tones which were more similar to each other in their pitch in the test also produced more correct intonation contours, similar to their teachers. There were no significant results for other factors, suggesting that learners' gender, lexical proficiency or musical experience are not related to the acquisition of L2 intonation in an advanced EFL classroom setting. We also reported no multicollinearity between the independent variables in the VIF-scores.

Since a traditional multiple linear regression model may not have uncovered all the underlying relations between the acquisition of near-native L2 English intonation and musical hearing or musical experience, we investigated systematically for chances of naturally good musical hearing skills or musical experience to contribute to the production of L2 intonation. Based on the results of the musical hearing tests and the musical experience survey, we categorized our participants into four types of L2 learners in a two-by-two matrix: 1) participants with good musical hearing test scores (i.e., having a result below 6 Hz in the *Adaptive Pitch Test* and/or above 70% in the *Tonedeaf Test*) and musical experience (i.e., singing and/or playing a musical instrument); 2) participants with good musical hearing test scores without musical experience; 3) participants with poor musical hearing test scores but with musical experience; and 4) participants with poor musical hearing test scores and without musical experience. The number of participants and their average results before and after training are provided in Table 7.

Table 7 Division of participants into musicians and non-musicians with good or poor musical hearing test scores (average intonation scores before (BT) and after training (AT) provided in brackets)

		Musical experience		Total
		Musicians	Non-musicians	
Musical hearing	Good	12 (.43 BT, .66 AT)	13 (.53 BT, .71 AT)	25 (.48 BT, .68 AT)
test scores	Poor	6 (.48 BT, .65 AT)	19 (.50 BT, .64 AT)	25 (.49 BT, .64 AT)
Total		18 (.44 BT, .66 AT)	32 (.51 BT, .67 AT)	50 (.49 BT, .66 AT)

A two-way ANOVA was run on the sample of 50 participants to examine the effect of musical hearing test scores and musical experience on the production of correct intonation scores before training. Residuals followed a normal distribution ($\alpha = .05$, $p = .07$) without outliers ($k = 1.5$). Although we found no significant interaction between the effects ($F(1, 46) = .33$, $p = .74$), we observed that participants without musical experience were able to achieve higher results than active musicians before training ($F(1, 46) = 3.04$, $p = .08$). There was no significant difference between the 25 participants who scored higher and lower on the musical hearing tests ($F(1, 46) = .10$, $p = .74$). When observing the distribution of the results in Figure 7, we can see that non-musicians with good musical hearing test scores were able to achieve the highest scores before training. Interestingly, some participants with musical experience and good musical hearing test results scored lower than participants with poor musical hearing test scores or no musical experience.

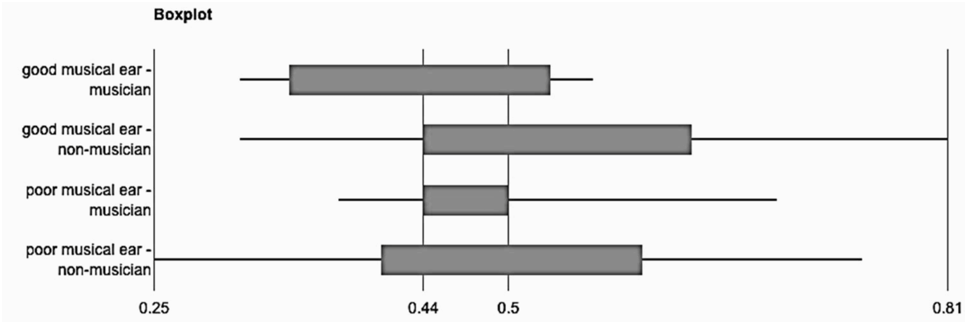


Figure 7 Intonation scores before training by learner type

We conducted a similar analysis on the same sample of 50 participants to examine the effect of musical hearing test scores and musical experience on the production of correct intonation scores after the training. Residuals followed a normal distribution ($\alpha = .05$, $p = .02$) without outliers ($k = 1.5$). We found no significant interaction between the effects ($F(1, 46) = .62$, $p = .43$) and there was no significant difference in the intonation scores between participants with good and poor musical hearing test scores ($F(1, 46) = .06$, $p = .79$) or between participants with and without musical experience ($F(1, 46) = .86$, $p = .35$). By observing the distribution of the results in Figure 8, we were able to determine that some non-musicians with good musical hearing test results were able to achieve higher intonation scores after training. By comparing these results with the intonation scores before training by learner type, we can see that the mean results for all learner types improved, but most of the participants with good musical hearing test scores produced more correct intonation patterns after training. These results suggest that participants with a good musical ear but no musical experience could have been positively stimulated during the accent training course and used their natural talent to their advantage in the acquisition of L2 intonation patterns.

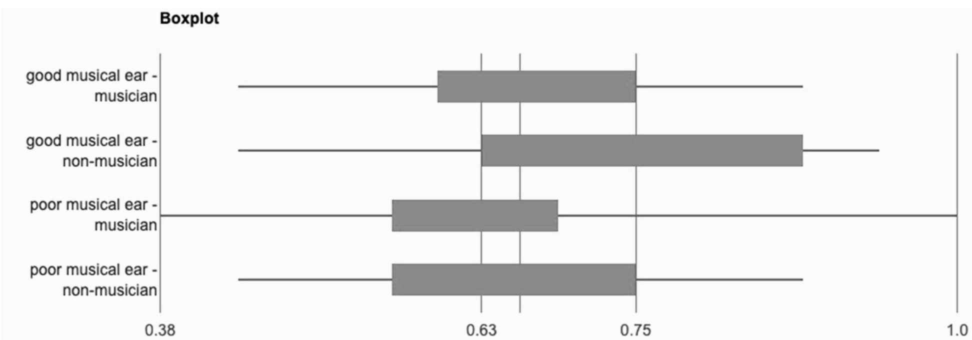


Figure 8 Intonation scores after training by learner type

6. Discussion

The results of the study are threefold. Our first research question asked whether Polish advanced learners of English were able to produce more native-like intonation patterns after training. We found a noticeable improvement in the production of correct GB intonational phrases after two academic semesters of an accent training course, combined with an English phonetics and phonology course. The courses provided the participants with practical pronunciation skills and phonetic awareness in order to acquire the specific features of the GB accent, and, consequently, produce similar intonational phrases to those of their accent training teachers. This finding shows that L2 intonation is both learnable (participants significantly improved their scores over time) and teachable (participants replicated their teachers' pronunciation) to a high level of proficiency in a formal learning environment. Interestingly, we also observed that not all intonational phrases were acquired with the same rate of success. The most difficult intonation patterns were the fall-rise, expressing non-finality, and the rise-fall, used for strong approvals. As these intonation patterns are relatively complex for L2 learners of English and not typically found in Polish speech, they are rarely used by Polish learners of English without formal accent training. At the same time, participants made a noticeable progress in *wh*-questions and tag questions. These results show that a change from a rising tone, used commonly in Polish questions (Mikoś, 1976), to the more typical falling pattern found in GB can be achieved by advanced learners of English after two semesters of formal accent training.

Our second research question considered participants who scored better in the musical hearing tests and whether they produced more correct intonation patterns after training. We found that participants who scored better in the *Adaptive Pitch Test* also produced more correct intonation patterns after training. That said, we found no significant relationship between participants' intonation scores and the *Tonedeaf Test* results. It should be noted that the results of both tests may differ from other musical hearing tests conducted on a wider population (see e.g., Barbaroux et al., 2020 for French non-musicians) as they usually differ across cultures and sample sizes. The tests used in this study were also recently used in other studies (Ning, 2020; Palomar-García et al., 2020), where Vietnamese learners of Mandarin scored comparably to the Polish speakers comprising our sample. Both of these studies have confirmed the validity of the tests and their relevance for studying the relationship between musical hearing and L2 language processing. The tests were used because their design aligned with the aims of this study, that is, researching the perception of tones and contrasting these cognitive skills with the ability of producing prosodic features in a second language. While the results of this study show that it is still possible for a learner with poor musical hearing test scores to produce native-like intonation contours after training, participants who scored higher in the *Adaptive Pitch Test* were able to correctly produce more of them. This finding reveals that being a good listener can be an asset in the production of more native-like L2 intonation patterns. Alternatively, it is possible to interpret these results as an indication that recognizing pitch change in the *Adaptive Pitch Test* is important for students practicing English intonation in the classroom, where rises and falls will be frequently used terms (see Zybert & Stępień, 2009).

Our final research question inquired whether participants with musical experience produced more correct intonation patterns after training, regardless of their musical hearing test results. We did not find any significant relationship between participants' musical experience and more accurate production of L2 intonation. Unlike former studies suggesting a strong link between musical practice and language skills (e.g., Chobert & Besson, 2013; Pastuszek-Lipińska, 2008), this result implies that musical background might not play a key role during a formal accent training course. Alternatively, it might suggest that the accent training course combined with a practical phonetics and phonology course could help all learners acquire L2 intonation and compensate for the lack of former musical practice. However, this does not negate the fact that superior pitch perception, even without formal music education, can be related to more accurate production of L2 intonation patterns.

In regard to long-term language acquisition – especially nowadays, when students are exposed to a great deal of audio and video material in native-spoken English – it is difficult to discern a single, overriding factor responsible for facilitating the acquisition of certain phonetic skills in learners. The results of our study suggest that formal phonetic instruction and practice, combined with finer pitch perception, can raise the success rate for learning foreign language intonation contours. In this study, we mentioned only the contours (i.e., falling tones, rising tones, etc.), but we realize that the acoustic signal in suprasegmental phonology involves not just F_0, but also pitch register, pitch span, rhythm, etc. Thus, it would be interesting to examine these parameters in future research. This study provides tentative evidence for how musical hearing can correlate to the acquisition of L2 intonation, using similar methodology to previous works by the authors investigating L2 vowel production (Jekiel & Malarski, 2021) and L2 rhythm (Jekiel, 2022). The current results refer to a rather narrow context of learning intonation in L2 English by Polish learners – whether they are applicable to other language pairs requires further research.

Despite the efforts put into designing a careful methodology and its longitudinal nature, there were several factors this study did not control for. First, the participants learning the GB accent were instructed by four different teachers.

Although the curriculum was the same for all groups, the impact of the individual differences in teaching styles the teachers may have presented could have been relatively strong. Another limitation of our study is the lack of control for motivation and other related variables, such as grit or general talent for language learning, apart from the musical context. These variables are widely discussed in the field of second language learning and teaching, and their role may have proven very interesting in discussing the results. Finally, the results are based on the data obtained from two recording sessions. An additional recording session after another academic year could have determined whether the gains were retained over time and point to the right direction showing whether the best-performing participants would have still been high-scoring in the delayed post-test.

7. Conclusions

The present study has demonstrated that superior musical hearing is correlated with more accurate production of L2 intonation patterns. Fifty Polish advanced learners of English were recorded reading a series of dialogues focusing on different intonational phrases before and after a two-semester accent training course, supplemented with an English phonetics and phonology course. The participants also completed two musical hearing tests assessing pitch perception and melody discrimination, and a musical experience survey. After comparing the participants' intonation patterns with the model provided by their accent training teachers, we compared their intonation scores with the musical hearing test results and survey responses. We found that superior pitch perception can be related to more native-like L2 intonation as participants with higher scores in the *Adaptive Pitch Test* also produced more accurate intonation contours, similar to those of their teachers. Although we observed that students with higher musical hearing test scores produced more correct L2 intonation patterns, students with poor musical hearing test scores and no musical experience also improved, suggesting that accent training in a formal classroom setting can lead to successful acquisition of L2 intonation regardless of students' musical hearing skills.

Acknowledgements

This research was supported by the National Science Center in Poland, Grant 2014/15/N/HS2/03865. Principal investigator: Mateusz Jekiel. Recipient: Adam Mickiewicz University, Poznań, Poland.

References

Alexander, J. A., Wang, P. C. M., & Bradlow, A. R. (2005). Lexical tone perception in musicians and non-musicians. *Proceedings of Interspeech 2005* (pp. 397-400). https://www.isca-speech.org/archive/pdfs/interspeech_2005/alexander05_interspeech.pdf

Aronsson, B. (2014). Prosody in the foreign language classroom - always present, rarely practised? *Journal of Linguistics and Language and Teaching, 5*, 207-225.

Barbaroux, M., Noreña, A., Rasamimanana, M., Castet, E., & Besson, M. (2020). From psychoacoustics to brain waves: A longitudinal approach to novel word learning. *Journal of Cognitive Neuroscience*, *33*(1), 8-27.

Beckman, M. E., & Elam, G. A. (1997). *Guidelines for ToBI labeling, version 3.0*. The Ohio State University.

Behroozmand, R., Ibrahim, N., Korzyukov, O., Robin, D. A., & Larson, C. R. (2014). Left-hemisphere activation is associated with enhanced vocal pitch error detection in musicians with absolute pitch. *Brain and Cognition, 84*(1), 97-108. https://doi.org/10.1016/j.bandc.2013.11.007

Bidelman, G. M., & Krishnan, A. (2009). Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem. *Journal of Neuroscience, 29*(42), 13165-13171. https://doi.org/10.1523/jneurosci.3900-09.2009

Boersma, P., & Weenink, D. (2022). Praat: Doing phonetics by computer [Computer program]. Version 6.3.03. Retrieved December 17, 2022 from http://www.praat.org/

Brugos, A., Shattuck-Hufnagel, S., & Veilleux, N. (2006). *Transcribing prosodic structure of spoken utterances with ToBI.* MIT Open Courseware. https://ocw.mit.edu/courses/6-911-transcribing-prosodic-structure-of-spoken-utterances-with-tobi-january-iap-2006/

Chapman, M. (2007). Theory and practice of teaching discourse intonation. *ELT Journal 61*(1), 3-11. https://doi.org/10.1093/elt/ccl039

Chen, A., Stevens, C. J., & Kager, R. (2017). Pitch perception in the first year of life, a comparison of lexical tones and musical pitch. *Frontiers in Psychology, 8*(297), 1-8. https://doi.org/10.3389/fpsyg.2017.00297

Chobert, J., & Besson, M. (2013). Musical expertise and second language learning. *Brain Sciences, 3*(2), 923-940. https://doi.org/10.3390/brainsci3020923

Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge University Press.

Dankovicová, J., House, J., Crooks, A., & Jones, K. (2007). The relationship between musical skills, music training, and intonation analysis skills. *Language and Speech, 50*(2), 177-225. https://doi.org/10.1177/00238309070500020201

Demirezen, M. (2009). An analysis of the problem-causing elements of intonation for Turkish teachers of English. *Procedia – Social and Behavioral Sciences*, *1*(1), 2776-2781. https://doi.org/10.1016/j.sbspro.2009.01.492

Derwing, T. M. (2008). Curriculum issues in teaching pronunciation to second language learners. In G. H. Edwards & M. L. Zampini (Eds.), *Phonology and second language acquisition* (pp. 347-369). John Benjamins.

Deutsch, D. (2002). The puzzle of absolute pitch. *Current Directions in Psychological Science, 11*(6), 200-204. https://doi.org/10.1111/1467-8721.00200

Draxler, J. Ch., & Jänsch, K. (2019). *SpeechRecorder – a universal platform independent multi-channel Audio Recording Software* (Version 4.4.50) [Computer software]. https://www.bas.uni-muenchen.de/forschung/Bas/software/speechrecorder/

Dufour, S., & Nguyen, N. (2013). How much imitation is there in a shadowing task? *Frontiers in Psychology, 4*, 1-7. https://doi.org/10.3389/fpsyg.2013.00346

Fenk-Oczlon, G., & Fenk, A. (2009). Some parallels between language and music from a cognitive and evolutionary perspective. *Musicae Scientiae, 13*(2), 201-226. https://doi.org/10.1177/1029864909013002101

Grabe, E., & Karpinski, M. (2003). Universal and language-specific aspects of intonation in English and Polish. *Proceedings of the 15th International Congress of Phonetic Sciences*, 1061-1064. https://www.internationalphoneticassociation.org/icphs/icphs2003

Gussmann, E. (2007). *The phonology of Polish*. Oxford University Press.

Gut, U. (2009). *Non-native speech: A corpus-based analysis of phonological and phonetic properties of L2 English and German*. Peter Lang.

He, X., van Heuven, V., & Gussenhoven, C. (2012). The selection of intonation contours by Chinese L2 speakers of Dutch: Orthographic closure vs. prosodic knowledge. *Second Language Research, 28*(3), 283-318. https://doi.org/10.1177/0267658312439668

Intartaglia, B., White-Schwoch, T., Kraus, N., & Schön, D. (2017). Music training enhances the automatic neural processing of foreign speech sounds. *Scientific Reports, 7*, 1-7. https://doi.org/10.1038/s41598-017-12575-1

Jekiel, M. (2022). L2 rhythm production and musical rhythm perception in advanced learners of English. *Poznan Studies in Contemporary Linguistics, 58*(2), 315-340. https://doi.org/10.1515/psicl-2022-0016

Jekiel, M., & Malarski, K. (2021). Musical hearing and musical experience in second language English vowel acquisition. *Journal of Speech, Language, and Hearing Research, 64*(5), 1666-1682. https://doi.org/10.1044/2021_JSLHR-19-00253

Jentschke, S., & Koelsch, S. (2009). Musical training modulates the development of syntax processing in children. *NeuroImage, 47*, 735-744. https://doi.org/10.1016/j.neuroimage.2009.04.090

Kang, O. (2013). Relative impact of pronunciation features on ratings of non-native speakers' oral proficiency. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference* (pp. 10-15). Iowa State University.

Kurt, B., Medlin, J., & Tessarolo, A. (2014). The perception of prosodically ambiguous intonation patterns by L2 English learners and the effects of instruction. *Concordia Working Papers in Applied Linguistics, 5*, 354-372.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods, 44*, 325-343. https://doi.org/10.3758/s13428-011-0146-0

Magne, C., Schön, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience, 18*(2), 199-211. https://doi.org/10.1162/jocn.2006.18.2.199

Mandell, J. (2009). *Music tests*. Jake Mandell, MD. http://jakemandell.com/

Marie, C., Delogu, F., Lampis, G., Belardinelli, M. O., & Besson, M. (2010). Influence of musical expertise on segmental and tonal processing in Mandarin Chinese. *Journal of Cognitive Neuroscience, 23*, 2701-2715. https://doi.org/10.1162/jocn.2010.21585

Mennen, I. (2004). Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics, 32*, 543-563. https://doi.org/10.1016/j.wocn.2004.02.002

Mikoś, M. J. (1976). Intonation of questions in Polish. *Journal of Phonetics*, *4*(3), 247-253.

Milovanov, R., Pietilä, P., Tervaniemi, M., & Esquef, P. A. (2010). Foreign language pronunciation skills and musical aptitude: A study of Finnish adults with higher education. *Learning and Individual Differences, 20*, 56-60. https://doi.org/10.1016/j.lindif.2009.11.003

Mitterer, H., & Müsseler, J. (2013). Regional accent variation in the shadowing task: Evidence for a loose perception–action coupling in speech. *Attention, Perception, & Psychophysics*, *75*, 557-575. https://doi.org/10.3758/s13414-012-0407-8

Moreno, S., & Besson, M. (2006). Influence of musical training on pitch processing: Event-related brain potential studies of adults and children. *Annals of the New York Academy of Sciences, 1060*, 93-97. https://doi.org/10.1196/annals.1360.054

Ning, L. (2020). Musical memory and pitch discrimination abilities as correlates of vocal pitch control for speakers with different tone and musical experiences. *Proceedings of the 10th International Conference on Speech Prosody 2020*, 611-615, https://doi.org/10.21437/SpeechProsody.2020-125

Nooteboom, S. (1997). The prosody of speech: Melody and rhythm. In J. Laver & W. J. Hardcastle (Eds.), *The handbook of phonetic sciences* (pp. 640-673). Blackwell Publishers.

Ott, C. G. M., & and Jäncke, L. (2013). Processing of self-initiated speech-sounds is different in musicians. *Frontiers in Human Neuroscience*, *7*, 1-15. https://doi.org/10.3389/fnhum.2013.0004

Palomar-García, M. Á., Olcina-Sempere, G., Hernández, M., Parcet, M. A., Mandell, J., & Ávila, C. (2020). The Jake Mandell test as a measure of individual differences in pitch discrimination: Validity and reliability properties. *Revista Electrónica Complutense de Investigación en Educación Musical, 17*, 143-151. https://doi.org/10.5209/reciem.66326

Pastuszek-Lipińska, B. (2008). Musicians outperform nonmusicians in speech imitation. *Lecture Notes in Computer Science, 4969*, 56-73. https://doi.org/10.1007/978-3-540-85035-9_4

Patel, A. D. (2008). *Music, language, and the brain*. Oxford University Press.

Patel, A. D., Foxton, J. M., & Griffiths, T. D. (2005). Musically tone-deaf individuals have difficulty discriminating intonation contours extracted from speech. *Brain and Cognition, 59*, 310-313. https://doi.org/10.1016/j.bandc.2004.10.003

Pijper, J. R. de. (1983). *Modelling British English intonation*. Foris Publications.

Roach, P. (2000). *English phonetics and phonology: A practical course*. (3rd ed.) Cambridge University Press.

Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology, 41*(3), 341-349. https://doi.org/10.1111/1469-8986.00172.x

Schön, D., Gordon, R., Campagne, A., Magne, C., Astésano, C., Anton, J., & Besson, M. (2010). Similar cerebral networks in language, music and song perception. *NeuroImage, 51*, 450-461. https://doi.org/10.1016/j.neuroimage.2010.02.023

Seashore, C. E., Lewis, D., & Saetveit, J. (1960). *Seashore measures of musical talents*. Psychological Corporation.

Slevc, L., & Miyake, A. (2006). Individual differences in second language proficiency: Does musical ability matter?. *Psychological Science, 17*, 675-681. https://doi.org/10.1111%2Fj.1467-9280.2006.01765.x

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*, 188-196. https://doi.org/10.1016/j.lindif.2010.02.004

Weckwerth, J, Janicka, K., Dziubalska-Kołaczyk K., Scheuer S., Wolfram-Romanowska, D., & Kul, M. (2006). Native standards or non-native ELF: Which

English to teach in the 21st century? In K. Dziubalska-Kołaczyk (Ed.), *IFAtuation: A life in IFA. A festschrift for Professor Jacek Fisiak on the occasion of his 70th birthday* (pp. 235-259). Adam Mickiewicz University Press.

Wells, J. C. (2014). *English intonation: An introduction*. Cambridge University Press.

Willems, N. (1982). *English intonation from a Dutch point of view*. Foris Publications.

Wong, P., & Perrachione, T. (2007). Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics, 28*(4), 565-585. https://doi.org/10.1017/S0142716407070312

Wong, P. C., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience, 10*(4), 420-422. https://doi.org/10.1038/nn1872

Zarate, J. M., Ritson, C. R., & Poeppel, D. (2012). Pitch interval discrimination and musical expertise: Is the semitone a perceptual boundary? *The Journal of the Acoustical Society of America, 132*, 984-993. https://doi.org/10.1121/1.4733535

Zatorre, R. J., & Baum, S. R. (2012). Musical melody and speech intonation: Singing a different tune. *PLoS Biology, 10*(7), 1-6. https://doi.org/10.1371/journal.pbio.1001372

Zybert, J., & Stępień, S. (2009). Musical intelligence and foreign language learning. *Research in Language, 7*, 99-111. https://doi.org/10.2478/v10015-009-0007-4

APPENDIX

Dialogues presented to participants (based on Wells, 2014)

Dialogue 1

A: What are your plans? Are you going to the concert tonight?
B: Well, not really. I'm going for a jog in the park.
A: Really? It'll rain in a minute!
B: I don't think so. Look, there's still some sun out there. But I'd better take my coat.
A: You'd better.

Dialogue 2

A: What are you drinking?
B: Coffee.
A: Neat! Let me have some.
B: Hands off my drink!
A: I only want to taste it...
B: You're broke again, aren't you?
A: Don't worry, I'll have some money soon.
B: In that case, here you go.

Dialogue 3

A: Hello, sir. How can I help you? Would you like another beer?
B: Thanks. The match is on Saturday, isn't it? (The customer isn't sure)
A: No, I think it was pushed a day ahead.
B: Oh no, so the match is on Sunday, isn't it? (The customer is now sure)
A: You will have to check that.
B: How did you like the football match yesterday?
A: Actually, I don't really watch football.

Dialogue 4

A: John! This is your frog, isn't it?
B: It's not a frog. It's a toad.
A: You know where I found it, don't you?
B: Oh no, was it in the pocket of your jacket again? Is the jacket all right?
A: I've checked it and it's okay. It's a run-of-the-mill jacket anyway.
B: Shall I pay for the cleaning?
A: Actually, let's talk about your homework. Did you finish the essay?
B: I've finished the introduction...