languages

MDPI

Article

Parallels Between Second Language Mastery and Musical Proficiency: Individual Differences in Auditory Phonological Pattern Recognition

Markus Christiner [1,2,*] and **Christine Groß** [2,3,4]

1 Department of Psychology, University of Graz, 8010 Graz, Austria
2 Jazeps Vitols Latvian Academy of Music, LV-1050 Riga, Latvia; christine.michaela.gross@jvlma.lv
3 Department of Child and Adolescent Psychiatry, Saarland University, 66421 Homburg, Germany
4 School of Health, Education, and Social Sciences, SRH University Heidelberg, 69123 Heidelberg, Germany
* Correspondence: markus.christiner@uni-graz.at

Abstract

Research has shown that language ability can vary enormously depending on variables such as musical ability, musical training, and second and/or foreign language experience. In this study, we simulated initial foreign language learning conditions in which learners must recognize and match unfamiliar language input. We recruited 500 participants with different levels of foreign language experience, different levels of musical training and different socio-economic backgrounds. Their auditory phonological pattern recognition ability, short-term memory (STM) capacity, musical ability, musical self-estimation, educational status, and socio-economic status (SES) were assessed. Both overall and group-specific analyses were conducted to investigate the impact of these variables. For the group-specific analysis, participants were assigned to four groups based on the presence or absence of musical training and extensive foreign language experience. For the overall analysis, regression models were applied to the entire sample to examine the combined effects of all variables. Group-specific analyses revealed that both musical training and extensive foreign language experience contributed to individual differences in the ability to recognize phonological patterns in unintelligible auditory stimuli. A key finding was that musical training appeared to have a stronger influence on auditory phonological pattern recognition than extensive foreign language experience, particularly in the early stages of language learning. This suggests that musical training may exert a greater impact on initial phonetic acquisition processes than extensive foreign language proficiency, especially when the language stimuli are relatively poor in linguistic content. The overall analysis revealed that musical variables, short-term memory capacity, socioeconomic status, and educational status all contributed to individual differences in auditory phonological pattern recognition. Notably, the most significant finding of the overall analysis was the association between SES and auditory phonological pattern recognition in unfamiliar speech—a result that challenges the notion of aptitude measures as stable and environment-independent and highlights the potential influence of environmental factors on this capacity.

Keywords: perception ability; musicality; musicians; foreign language experience; socio-economic status; aptitude

check for updates

Academic Editor: Adrian P. Simpson

Received: 25 June 2025
Revised: 13 October 2025
Accepted: 23 October 2025
Published: 27 October 2025

Citation: Christiner, M., & Groß, C. (2025). Parallels Between Second Language Mastery and Musical Proficiency: Individual Differences in Auditory Phonological Pattern Recognition. *Languages*, *10*(11), 272. https://doi.org/10.3390/languages10110272

1. Introduction

Research on individual differences in foreign language acquisition has focused on multiple domains that explain individual differences in foreign language capacity at different stages of learning, for example: specific learning styles/learning strategies (O'Malley & Chamot, 1990), age (Singleton & Lengyel, 1995), intelligence quotient (Robinson, 2019), personality factors (Dörnyei, 2006), anxiety (Baran-Łucarz, 2016; M. G. Calvo & Eysenck, 1996; Onwuegbuzie et al., 1999), attitudes (Gardner, 1990), motivation (Dörnyei et al., 2015), interest (Liu, 2022), socio-economic status (SES), bi- and multilingualism (Christiner et al., 2021; Dewaele & Wei, 2012; Kaushanskaya & Marian, 2009; Korecky-Kröll et al., 2019; Singleton & Lengyel, 1995; Wrembel, 2015), overlaps between music and language (Christiner, 2018; Christiner et al., 2023, 2022a, 2022c; Delogu & Zheng, 2020), language aptitude and ability (Skehan, 2012, 2019; Wen, 2012; Wen et al., 2017; Wen & Skehan, 2011), short-term memory (Coumel et al., 2023; Gathercole et al., 1997) and working memory ability (Archibald, 2017; Baddeley, 2003; Kofler et al., 2020; Linck et al., 2013; Loosli et al., 2012). Therefore, individual characteristics, affective factors, linguistic background, socio-economic status, and cognitive abilities emerge as pivotal areas of focus in foreign language research, given their significant impact on language acquisition and proficiency. Among these domains, age is highlighted as a key variable influencing language capacity at different learning stages.

In general, it is accepted that age is a critical factor in the acquisition of foreign languages, indicating that the younger the age additional languages are acquired, the better the achievement will be (Johnson & Newport, 1989; Lenneberg, 1967). As a consequence, foreign language learning becomes more difficult in older ages and affects particularly phonetic ability, a result of a natural tendency and preference towards the mother tongue (Christiner & Reiterer, 2013). Phonetic ability has also been proposed to be the sole capacity ultimately subject to a critical period (Moyer, 2014). Pioneering research in the field highlighted the rarity of adult non-native speakers achieving native-like foreign language pronunciation (Selinker, 1972). Exceptional abilities such as foreign language aptitude are domain-specific (Jørgensen, 2008) and usually found in the upper 10 percent of the normal distribution (Gagné, 2008). From a developmental perspective, aptitude is seen as the foundation upon which talent is built through practice. While some argue that high talent is largely based on innate ability that manifests more easily in younger ages (Gagné, 2004, 2008), others challenge this notion, emphasizing the crucial role of extensive practice and equal opportunities in achieving excellence (Howe et al., 1998). This perspective is supported by evidence indicating that achieving high excellence in a specific domain is closely tied to practice and learning experiences rather than predetermined genetic factors (Howe et al., 1998).

Research on language aptitude often involves tasks that incorporate unfamiliar language material for assessment. Researchers who studied adult exceptional foreign language achievers propose that their success is largely attributed to advanced perceptual, cognitive abilities (Doughty et al., 2008; Linck et al., 2013) and vocal skills (Christiner et al., 2023, 2022a, 2022c). This is also highly plausible, especially when considering the phonetic challenges encountered in foreign language learning during initial learning stages.

In general, it is widely acknowledged that adult learners encounter numerous challenges in acquiring foreign languages (Abrahamsson & Hyltenstam, 2009; DeKeyser et al., 2010; Han, 2013). For instance, foreign language learners tend to use sounds which resemble those of their native language, resulting in perceptual and productive transfer errors when attempting the target language (Christiner, 2020; Patel, 2007; Yang et al., 2022). Depending on the language, foreign language learners can also be concerned with the learning of very specific language features that are not rooted in their mother tongue (Werker & Tees, 2005). It is also well-observed that foreign language learners find it challenging to comprehend utterances when spoken by a native speaker, although they probably understand their meaning when the utterance is presented visually. Foreign language learners often face challenges in identifying word boundaries in continuous speech, as they tend to apply segmentation strategies from their native language (Patel, 2007). The latter leads to perceptual difficulties and to problems in being able to categorize sounds into meaningful units. Consequently, the acquisition of phonetic structures emerges as a predominant factor in the early stages of learning foreign languages.

In the initial phase of learning a foreign language, being musically inclined or musically trained is helpful due to the shared cognitive processes involved in processing linguistic and musical stimuli (Besson et al., 2011; Peretz et al., 2015). Therefore, in the initial stages of foreign language learning, overlaps between musicality and language have been very frequently observed (Christiner et al., 2023, 2022b; Milovanov, 2009; Milovanov & Tervaniemi, 2011; Pastuszek-Lipinska, 2008). Thus, it comes as no surprise that over the past two decades several models on the overlap between music and language have been proposed. For instance, the shared sound category learning mechanism hypothesis suggests that the mechanism for learning sound categories show overlaps between music and language (Patel, 2007). This suggests that the influence of musical training on language abilities can be ascribed to the refinement of a fundamental, cross-domain sound learning mechanism. While extensive literature has demonstrated that musical skills can predict language ability, some studies have reported contrary findings or alternative explanations for their relationship.

For instance, singing as a sensory phenomenon has been linked to productive phonetic skills such as the pronunciation of foreign languages, the ability to mimic foreign accents, and the ability to memorize new vocabulary (Christiner et al., 2021, 2023, 2022a; Christiner & Reiterer, 2018; Franco et al., 2021; Ludke, 2018; Ludke et al., 2014; Groß et al., 2022). The overlaps between singing and foreign language pronunciation may be mediated by enhanced sensorimotor ability (Christiner et al., 2022a, 2022b, 2022c).

Perceptual research has shown that musicians possess enhanced auditory skills, enabling precise detection of nuances in pitch, tone, and rhythm (Besson et al., 2011; François et al., 2013; Kraus & Chandrasekaran, 2010; Thompson et al., 2004). Heightened auditory perception aligns with enhanced capabilities in tone languages discrimination (Christiner et al., 2022b), speech segmentation (François et al., 2013) and perception of non-native tones (Alexander et al., 2005), or higher language processes such as reading comprehension (Seither-Preisler et al., 2014; Partanen et al., 2022; Eccles et al., 2021).

Some studies report conflicting findings on the link between musical capacity and language abilities. For instance, although musical training is associated with language-related skills, this relationship may weaken or disappear once general cognitive abilities are considered (Swaminathan et al., 2018). Such results suggest alternative explanations, indicating that observed associations may largely reflect underlying cognitive and perceptual skills common to both domains (Swaminathan & Schellenberg, 2020; Kragness et al., 2021). Other research highlights that specific musical skills—especially singing ability—are more closely related to language functions such as pronunciation, while associations with other musical skills are less consistent (Christiner et al., 2022a). Ultimately, the presence or absence of positive effects depends on multiple mechanisms and must be carefully examined.

People who speak two or more languages, or who acquire several languages during their lives, have a similar status to musicians from a linguistic point of view; they have a clear advantage when learning new languages (Gresele et al., 2013; Papagno & Vallar, 1995). Researchers distinguish between simultaneous and successive bilinguals/multilinguals. The latter possess improved auditory language skills (Gresele et al., 2013) and also show enhanced ability to acquire novel phonological forms (Hell & Mahn, 1997; Kaushanskaya

& Marian, 2009; Papagno & Vallar, 1995). This is dependent on several factors, two of which are pivotal for our study. In language research, phonological short-term memory has been defined as one of the most important predictors that explain individual differences in foreign language ability (Wen & Skehan, 2011). STM capacity is closely linked to the capacity to recall larger phonological units and stands out as a key cognitive factor predictive of advanced language skills among multilinguals (Baddeley et al., 1998). In addition, there are also training related benefits. It has been suggested that individuals with extensive foreign language capacity benefit from more refined and flexible long-term representations of phonological patterns, such as the ability to process and remember longer sequences of speech sounds (e.g., extended syllable strings or nonwords), which may facilitate faster acquisition of additional languages (Hell & Mahn, 1997; Kaushanskaya & Marian, 2009; Papagno & Vallar, 1995).

Previous research has demonstrated a significant association between initial foreign language aptitude measures and the educational backgrounds of participants' parents (Christiner, 2020), highlighting parental education as a key socioeconomic (SES) variable in the study of language aptitude. Despite this, research specifically examining the relationship between language aptitude and SES remains limited.

Socioeconomic status is a multifaceted construct that encompasses access to economic and social resources as well as social positioning, and is closely linked to educational and academic achievement (Hauser & Warren, 1997; Li et al., 2019; Szabó et al., 2019). Individual differences in language capacity have also been associated with SES. For example, higher SES among mothers and caregivers has been shown to positively predict language skills in children (Huttenlocher et al., 2010), while individuals from lower SES backgrounds tend to perform more poorly on language tasks (A. Calvo & Bialystok, 2014).

Research on foreign language acquisition in children further suggests that SES-related disparities extend beyond the quantity of language input to include differences in the quality of linguistic interactions and opportunities for language learning (Cartmill et al., 2013; Rowe, 2018). Importantly, the impact of socioeconomic disparities on language skills may persist into adulthood and continue to influence academic achievement (Hoff, 2013). These findings underscore the importance of investigating whether socioeconomic status similarly affects the ability to perceive and retain unfamiliar languages.

In previous research we noted that extensive foreign language training and musical status and musical training variables (e.g., familiar song singing) were good predictors for the ability to pronounce unfamiliar languages in adult participants, regardless of the languages which were used (Christiner, 2020; Christiner et al., 2021, 2022b, 2023). In this study, our focus shifted to examining whether we could replicate these findings using auditory phonological pattern recognition measures. Some language aptitude measures (e.g., LLAMA D; Meara, 2005) are based on the premise that the ability to identify brief phonological sequences in an unfamiliar language facilitates subsequent vocabulary acquisition. In the same vein, musicality assessments operate on analogous principles, requiring listeners to compare and judge previously heard patterns (e.g., short melodies) as identical or different. On this basis, both types of tasks recruit similar cognitive mechanisms for detecting and representing sequential auditory information. In this study, we addressed these specific overlaps and formulated the following research questions: Does the relationship between musical training and musical ability also extend to auditory phonological pattern recognition (Q1)? How are short-term memory, socioeconomic status (SES), educational level, and self-estimated musical ability related to the ability to recognize phonological patterns in unintelligible auditory stimuli (Q2)? Additionally, we aimed to investigate whether extensive foreign language training and musical background served as equally strong predictors of our variables under consideration (including scores on the Advanced

Measures of Music Audiation (AMMA) test, SES status, STM, educational level, and musical ability self-estimation), with particular emphasis on the ability to recognize phonological patterns in unintelligible auditory stimuli. We aimed to determine the extent to which extensive foreign language training and musical status distinguish adult participants in their performance on auditory phonological pattern recognition and related cognitive measures (Q3). To achieve this, we recruited participants both with and without extensive foreign language training, as well as individuals with and without musical training.

2. Materials and Methods

2.1. Participants

We recruited participants through online platforms, schools, and universities, providing them with information about the eligibility criteria they needed to meet. First, the most important criterion was that participants should neither speak, comprehend nor have studied Tagalog, Mandarin, Farsi and Japanese as a second, or foreign language. This was meant to simulate a foreign language learning setting in initial learning stages when individuals listen to unfamiliar utterances that they retain for a short period to ensure that participants had no prior knowledge about any of the languages included in the research design. The recruitment description also stipulated that we look for individuals with and without musical training without distinction of the genre they play or perform (including both non-professional and professional musicians), and individuals with and without extensive foreign language experience. The participants had no hearing or cognitive impairments. Informed consent was obtained from all subjects involved in the study. The ethical committee has granted approval for this study GZ. 39/106/63 ex 2024/25.

In order to be able to categorize the participants according to our criteria, they had to fill out a detailed questionnaire about their language, musical and educational background. The participants were categorized as participants with or without formal musical training according to the following criteria. Participants with no musical training had received no formal instruction on any musical instrument prior to the test sessions. Learning the recorder in school was not treated as musical training. Additionally, participants who had not received musical training were asked to indicate if they had learned to play a musical instrument informally. Participants were categorized as individuals with musical training by fulfilling two criteria. First, they had to have played a musical instrument for longer than 3 years on a regular basis before testing. Second, they had to regularly participate in musical activities at the time they were tested. Musicians who no longer actively practiced or played an instrument—sometimes referred to as “sleeping musicians”—were not included in this study. Since some participants stopped playing at different times in the past, establishing consistent and meaningful criteria for inclusion proved difficult. Participants falling into this category were excluded from the analysis. The questionnaire used for categorization has been provided in the Supplement (see Supplement Section S1).

Furthermore, the participants were differentiated based on their foreign language proficiency. All adult testees had undergone foreign language training in the course of their lives. Participants were categorized as adults with extensive foreign language training if they had demonstrated proficiency in speaking and comprehending at least 2 foreign languages at a high level. The participants were considered to have extensive foreign language proficiency if they were able to master at least two foreign languages, in addition to their mother tongue, at the B2 or C1 level. The foreign languages spoken included English, German, French, Spanish, Italian, Dutch, Russian, Croatian, Hungarian, Norwegian, Portuguese, Swedish, Slovak, Czech, and Finnish. We did not include participants who were simultaneous bilinguals but only successive bi-and multilinguals. The questionnaire is available in the Supplement Section (see Supplement Section S2).

According to the above description, we established 4 different groups: adults without musical training and without extensive foreign language training (NoMT + NoEFL; *N* = 91), adults without musical training but with extensive foreign language experience (NoMT + EFL; *N* = 125); adults with musical training but without extensive foreign language training (MT + NoEFL; *N* = 102), and adults with musical training and extensive foreign language experience (MT + EFL; *N* = 182). Given the number of variables defined, we have included a list of abbreviations to aid comprehension (see Table S2 in the Supplement).

A total of 500 participants were involved in the study, comprising 156 males and 344 females, with a mean age of 29.60 years (SD = 11.01). We also collected information about their native language and the educational status of the participants as well as their parents' educational backgrounds. Of the participants, 451 reported German as their native language, constituting the vast majority of the sample. The remaining participants were native speakers of Latvian (17), Croatian (8), English (7), Slovenian (5), Russian (4), and Bosnian (3). Additionally, Bulgarian and Polish each accounted for 2 native speakers, and Czech for 1 native speaker. The determination of educational status in this study adheres primarily to the guidelines outlined by the European Qualifications Framework (EQF). Educational status is represented on a scale ranging from 0 to 9, where 0 denotes 'no school education' and 9 signifies attainment of a 'doctoral degree' (for details see Supplement Section S3). The educational score of the participants is revered to as educational level of the participants in the analysis section.

The educational status of parents is frequently used as a criterion for defining socioeconomic status (SES) with the primary caregiver typically playing a more significant role. To account for the differential influence of the main caregiver's education, we applied a weighted approach, with the main caregiver's educational status assigned a higher weight (SES score = 0.75 $\times$ main parent/caretaker education + 0.25 s parent).

2.2. Auditory Phonological Pattern Recognition Measure

Auditory perception entails linking incoming sounds with previous experiences and adapting to environmental conditions and the anticipated relevance of the stimuli. Depending on the cognitive demands, it may involve processes such as sound classification, recognition, and identification, applicable to both speech and non-speech auditory signals (Scharine et al., 2009).

Within this broad framework, we developed a tool to assess auditory phonological pattern recognition—structurally similar to measures used in previous research (Christiner, 2020; Christiner et al., 2022a; Coumel et al., 2023). Auditory phonological pattern recognition is defined here as the ability to detect, segment, and match unfamiliar multi-syllabic speech sequences based on their phonological and prosodic patterns—independent of semantic or lexical knowledge. It evaluates individual differences by emphasizing holistic pattern recognition over the discrimination of specific phonological contrasts.

This approach aligns conceptually with established language aptitude paradigms, such as the LLAMA D test (Meara, 2005), which measures implicit learning of repeated auditory patterns, and musical aptitude assessments like Edwin Gordon's AMMA test (1989), which evaluate auditory pattern recognition through unfamiliar melodies. The LLAMA D test presents paired auditory sequences for comparison, mirroring the approach of musical aptitude tasks that match paired sequences. Our measure employs sequence recognition and is based on approaches similar to those used in LLAMA D and many musicality tasks, as all share the requirement that previously heard patterns must be compared and judged to be the same or different. Unlike classical speech perception tasks that emphasize phonetic-phonological contrasts (e.g., minimal pairs), this measure emphasizes the perception and retrieval of natural multi-syllabic speech units as integrated sequences. By incorporating multiple typologically diverse languages, it enhances reliability and generalizability. Analogous to how varied melodies enhance musical assessments, using different languages and speakers—each providing unique phonological sequences—enhances the reliability of our measure. Moreover, it parallels musicality measures, serving as an appendage focused on phonological sequences.

2.2.1. Stimulus Selection and Characteristics

Stimuli are drawn from the narrative The North Wind and the Sun, presented in four typologically diverse languages that are unfamiliar to the participants of this study, who were primarily native German speakers: Mandarin, Japanese, Farsi, and Tagalog. While each language has characteristic phonological features, these specific contrasts are not the focus of the test. The choice of diverse languages ensures participants' lack of prior familiarity and introduces natural phonetic and prosodic variability. The stimuli were produced by six native speakers per language to capture realistic variation in articulation and prosody.

2.2.2. Stimulus Preparation and Structure

For each language and speaker, approximately 24 samples—consisting of words and short phrases ranging from 3 to 11 syllables—were extracted at naturally occurring speech pauses and normalized for loudness to maintain comparability across recordings. Each trial consists of two parts: a string and a response stimulus. The string comprises eight segments per participant: four segments containing between 3 and 6 syllables and four segments containing between 7 and 11 syllables, sampling a spectrum of phonological complexity and prosodic length. Segments within the string are separated by brief silent intervals of 50 ms. After the string, a 2-s delay indicated by a change in screen color precedes the response stimulus.

Response stimuli consist of one to three units, which participants must identify as belonging to the preceding string. When multiple units are presented in a response, all must have appeared somewhere in the string for the response to be judged correct, regardless of order. For example, a response with three units is considered correct if all three appeared somewhere in the first part of the trial (the presentation string), regardless of their order in the response. All stimuli within a single trial are produced by the same native speaker, ensuring uniformity of speaker-specific prosodic cues. An example from Mandarin illustrating the specific linguistic units extracted for the study is provided in the Supplemental Materials (see Supplemental Section S4 Table S1).

In sum the task consists of 20 items and for each correct answer 0.05 points were given. Therefore, the maximum score someone can reach is 1 if all answers are correct.

2.3. Short-Term Memory Measure

The phonological short-term memory (STM) of the participants was assessed using forward and backward Wechsler digit span tasks (Tewes, 1994). In the forward task, participants were instructed to repeat a sequentially increasing series of numbers presented auditorily, while in the backward task, they were instructed to repeat the numbers in reverse order. The forward span task consisted of sequences ranging from three to nine items, whereas the backward span task included sequences ranging from two to eight numbers. Each participant received two scores: one for the forward task and one for the backward task. Scores were based on the number of items correctly repeated, with a maximum possible score of 14. In this study we used a composite score of both measures for the analysis (STM total).

2.4. Music Perception Ability

The AMMA test (Advanced Measures of Music Audiation, Gordon, 1989) was utilized to evaluate participants' music perception abilities. This test presented participants with pairs of musical statements, prompting them to identify whether these pairs were identical or exhibited rhythmic or tonal differences. Rhythmic variations included changes in tempo, meter, or duration, while tonal differences involved discrepancies in a few notes between the two statements. The test comprised 33 items, with participants completing three practice trials followed by 30 experimental trials. Of these, 10 pairs were identical, 10 differed rhythmically, and 10 differed tonally. The order of presentation for the trials was randomized. For the analysis in this study, we used the composite score (AMMA total) that consists of the results of all trials.

2.5. Statistical Analyses

The statistical analysis is divided into different steps. First correlational analysis was performed for the variables under consideration. Then, a regression analysis was conducted to address two primary research questions (Q1 and Q2): the first question focused on whether we could replicate previous findings using auditory phonological pattern recognition measures and their relationship with musicality variables and the second question explored further dimensions by investigating the relationships between STM, SES, educational level, self-estimation in musical ability, and the ability to discriminate unfamiliar languages. In the regression model, auditory phonological pattern recognition was the dependent variable.

To investigate the third research question (Q3), we conducted group analyses. The main aim was to examine whether extensive foreign language training and musical background serve as equally strong predictors of our variables under consideration, with particular interest in auditory phonological pattern recognition. We ran a MANOVA to reveal whether the variables AMMA total, SES status, STM, educational level, musical ability self-estimation, and auditory phonological pattern recognition differ in their mean values across the four groups (with/without extensive foreign language training, with/without musical training). As a follow-up analysis, we performed a discriminant analysis to illustrate how well these variables predict the group membership of our participants. Additionally, we conducted a separate ANOVA specifically focused on auditory phonological pattern recognition to allow for more detailed differentiation.

As gender ultimately did not influence perceptual language measures in previous research (Christiner, 2020; Coumel et al., 2023; Groß et al., 2023), gender differences were not part of our study. For transparency we include here a *t*-test which revealed that there was no significant difference between female (M = 0.67, SE = 0.009) and male participants (M = 0.67, SE = 0.007); $t(498) = 0.09$, $p = 0.93$. The descriptive statistics for the variables under consideration are provided in Supplementary Section S5, Table S4.

3. Results

3.1. Correlational Analysis

To assess the associations between the auditory phonological pattern recognition and the variables under consideration we performed correlational analysis. Table 1 shows the simple associations between the main variables under consideration.

Table 1. Simple associations of the language and musical variables.

Variable	AMMA Total	STM Total	Musical Self-Estimation	SES Status	Educational Level
Auditory phonological pattern recognition	0.435 **	0.303 **	0.174 **	0.248 **	0.216 **
AMMA total		0.285 **	0.146 **	0.142 **	0.133 **
STM total			0.057	0.182 **	0.105 *
Musical ability self-estimation				0.149 **	0.138 **
SES status					0.240 **

Note. STM = short-term memory. SES status = socio-economic status. * $p < 0.05$ (uncorrected, two-tailed). ** $p < 0.001$ (uncorrected, two-tailed).

3.2. Regression Analysis

A stepwise multiple linear regression analysis was conducted to examine the relationships between various factors and auditory phonological pattern recognition scores. To determine which variables to include in the regression model, we first performed correlation analyses (descriptives are provided in the Supplement, see Table S4). Variables that showed significant correlations with the auditory phonological pattern recognition score ($p < 0.05$) were selected as potential predictors for the regression model.

The results of the multiple regression analysis (Table 2) demonstrate that several factors contribute significantly to the variance in auditory phonological pattern recognition. Model 1 shows that musical perception (AMMA total) explains 19% of the variance ($R^2 = 0.19$, $p < 0.01$). With the addition of SES status in Model 2, the explained variance increases to 23% ($R^2 = 0.23$, $p < 0.01$). Model 3 incorporates STM total, further increasing the explained variance to 25% ($R^2 = 0.25$, $p < 0.01$). Educational level in Model 4 adds a small but significant contribution, raising the variance explained to 26% ($R^2 = 0.26$, $p < 0.01$). Finally, Model 5, which includes musical ability self-estimation, accounts for 27% of the variance in auditory phonological pattern recognition ($R^2 = 0.27$, $p < 0.05$). For AMMA total, SES status, STM total, educational level, and musical ability self-estimation, VIF values ranged from 1.000 to 1.122, suggesting that multicollinearity was not a concern in the present analyses. Residual diagnostics indicated that standardized residuals ranged from −2.73 to 2.59 (M = 0.00, SD = 0.99), suggesting no extreme outliers. The residuals were all within the expected range of ±3, supporting the assumption of normality and the adequacy of the regression model.

Table 2. Multiple Regression Models Explaining the Variance in Auditory Phonological Pattern Recognition.

		R	R^2	F Change	Sig. F Change	*B*	*SE B*	β	*p*
Model 1		0.44	0.19	116.44	0.000				
	Constant					0.35	0.03		
	AMMA total					0.01	0.00	0.44	<0.01
Model 2		0.47	0.23	22.78	0.000				
	Constant					0.31	0.03		
	AMMA total					0.01	0.00	0.41	<0.01
	SES status					0.01	0.00	0.19	<0.01

Table 2. *Cont.*

		R	R^2	F Change	Sig. F Change	B	$SE\ B$	β	p
Model 3		0.50	0.25	16.97	0.000				
	Constant					0.23	0.04		
	AMMA total					0.01	0.00	0.36	<0.01
	SES status					0.01	0.00	0.17	<0.01
	STM total					0.01	0.00	0.17	<0.01
Model 4		0.51	0.26	8.78	0.003				
	AMMA total					0.01	0.00		
	SES status					0.01	0.00	0.14	<0.01
	STM total					0.01	0.00	0.16	<0.01
	Educational level					0.01	0.00	0.12	<0.01
Model 5		0.52	0.27	4.10	0.044				
	AMMA total					0.00	0.00		
	SES status					0.01	0.00	0.13	<0.01
	STM total					0.01	0.00	0.16	<0.01
	Educational level					0.01	0.00	0.11	<0.01
	Musical ability self-estimation					0.01	0.00	0.08	<0.05

Dependent variable: Auditory Phonological Pattern Recognition.

3.3. MANOVA, Discriminant Analysis and ANOVA

We conducted a MANOVA to investigate potential differences in the mean values of the variables auditory phonological pattern recognition, AMMA total, SES status, STM, educational level, and self-estimated musical ability across groups. Table 3 presents the mean values of the variables under consideration. Using Pillai's trace there was a significant effect for group membership (NoMT + NoEFL, NoMT + EFL, MT + NoEFL, and MT + EFL) for the auditory phonological pattern recognition, AMMA total, SES status, STM, educational level, and self-estimated musical ability $V = 0.98$, $F(6{,}491) = 11.75$, $p < 0.001$.

The MANOVA was followed by a discriminant analysis, which revealed two discriminant functions (see Figure 1). The first explained 85.1% of the variance, canonical $R^2 = 0.31$, whereas the second explained only 12.9%, canonical $R^2 = 0.06$. The third function was non-significant and explained only 1.7% of the variance. In combination, these discriminant functions discriminated the groups, $\Lambda = 0.65$, $\chi2(18)$ 215.27, $p < 0.001$. $\Lambda = 0.9$, $\chi2(10) = 36.07$, $p = 0.001$, while the third function did not discriminate the groups $\Lambda = 0.9$, $\chi2(4) = 4.35$, $p = 0.361$. We used the recommended cutoff of value of 0.4 (Hair et al., 2014) to decide which of the standardized discriminant coefficients are large. The correlations between the outcomes and the first discriminant functions revealed that the loads onto the first function are large for the AMMA total ($r = 0.69$), the auditory phonological pattern recognition variable ($r = 0.64$), and the musical ability self-estimation score ($r = 0.53$). Although the STM score also correlated with the first function, it did not reach the upper limit to be considered a significant score ($r = 0.33$). Table 3 illustrates that the means of the variables were higher for groups with musical training compared to those without. The correlations between the outcomes and the second discriminant function were high for participants' educational level ($r = 0.62$) and SES ($r = 0.56$). Similarly, the means were higher for groups with extensive foreign language training compared to those without. To provide information about the significance of group differences in the main variable of interest, auditory phonological pattern recognition, we conducted an ANOVA. There was a significant difference in auditory phonological pattern recognition across the four groups, $F(3, 496) = 30.713$, $p < 0.001$, $\omega^2 = 0.15$, indicating a large effect size. Furthermore, a significant linear trend was observed, $F(1, 496) = 92.015$, $p < 0.001$, $\omega^2 = 0.16$, suggesting a progressive increase in auditory phonological pattern recognition across the groups in the order of NoMT + NoEFL, NoMT + EFL, MT + NoEFL, and MT + EFL, with a large effect size. Gabriel-corrected post-hoc analyses comparing the four groups were also performed to provide detailed insights into these differences. The results are available in the Supplemental Material (Table S3).

Table 3. Mean Values of Each Variable for the Four Groups.

Variables	Groups	Mean (*M*)	Standard Deviation (*SD*)
AMMA total	NoMT + NoEFL	47.73	8.57
	NoMT + EFL	47.82	7.08
	MT + NoEFL	54.76	7.20
	MT + EFL	55.20	7.99
STM total	NoMT + NoEFL	13.51	2.26
	NoMT + EFL	13.83	2.46
	MT + NoEFL	14.66	2.47
	MT + EFL	14.74	2.20
Musical ability self-estimation	NoMT + NoEFL	2.41	0.98
	NoMT + EFL	2.55	1.06
	MT + NoEFL	3.25	1.14
	MT + EFL	3.21	1.01
Auditory phonological pattern recognition	NoMT + NoEFL	0.58	0.13
	NoMT + EFL	0.64	0.10
	MT + NoEFL	0.69	0.11
	MT + EFL	0.71	0.11
SES status	NoMT + NoEFL	3.73	1.66
	NoMT + EFL	4.62	1.98
	MT + NoEFL	4.55	1.98
	MT + EFL	5.33	2.19
Educational level	NoMT + NoEFL	4.40	1.91
	NoMT + EFL	5.34	2.02
	MT + NoEFL	5.17	2.11
	MT + EFL	6.02	2.10

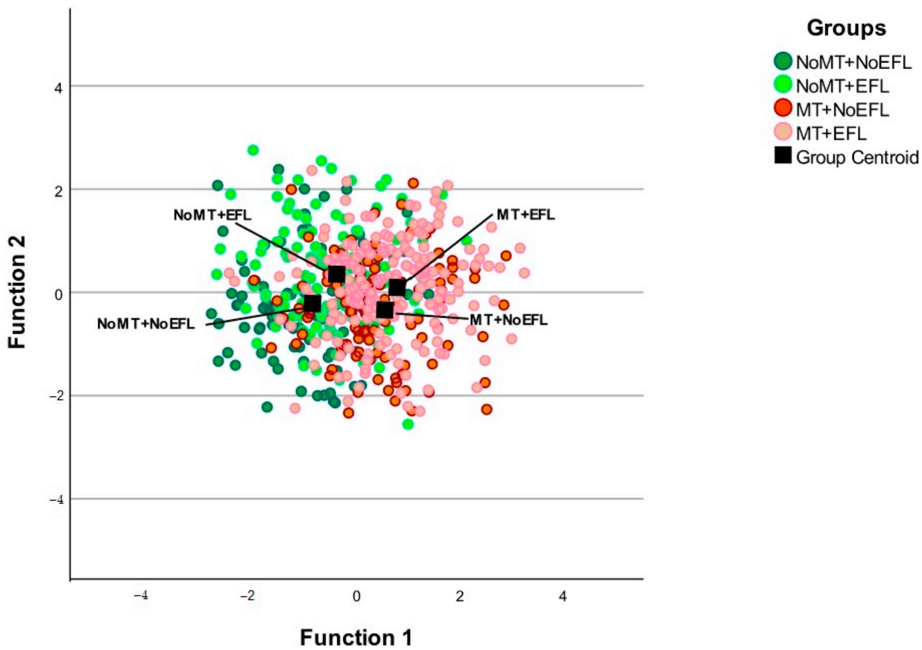

Figure 1. Function 1 primarily separates musically trained from untrained participants, showing high loadings on AMMA total scores, self-estimated musical ability, and auditory phonological pattern

recognition. In the plot, musically trained participants are shown in red and pink, whereas musically untrained participants are displayed in dark and light green. Function 2 differentiates participants with extensive foreign language training from those without, showing high loadings on education level and socioeconomic status (SES). The clusters display some minor overlap, which does not affect the interpretability of the discriminant structure or the observed group separation.

4. Discussion

In this study, our goal was to investigate whether there are connections between musical ability and auditory phonological pattern recognition, moving beyond previous research that primarily focused mostly on pronunciation tasks (Q1). In addition, we included additional variables such as short-term memory capacity, educational level, SES, and musical ability measures in the study to outline potential relationships to auditory phonological pattern recognition (Q2). Regression analysis revealed that musical variables (musical aptitude and self-estimation), STM capacity, educational status, and SES status predict individual differences in the ability to recognize auditory phonological patterns in unintelligible speech. Notably, the contributions of STM capacity and socioeconomic factors demonstrated significant effects on auditory phonological pattern recognition, highlighting their importance in understanding how various factors interact with musicality and auditory phonological pattern recognition.

Our analysis also focused on the relative influence of extensive foreign language experience and musical training on auditory phonological pattern recognition, along with other variables of this study (Q3). To address these overarching questions, we employed three main statistical approaches: group-based analyses using MANOVA, discriminant analysis and ANOVA. The results from the group-based analyses indicated that musically trained participants demonstrated not only enhanced musical ability but also improved auditory phonological pattern recognition, irrespective of whether they had received extensive foreign language training or not. We will not elaborate on the separation of groups with extensive foreign language experience from those without, based on their educational status and socioeconomic status (SES), as indicated in discriminant analysis function 2. This outcome aligned with our expectations and does not warrant further discussion. Based on the key findings of our study, the discussion will be structured into two parts. First, we will analyze the results of the regression analysis, focusing on the factors that significantly influence auditory phonological pattern recognition. Next, we will examine the findings from the group-based analyses, highlighting the crucial roles of musical ability and extensive foreign language experience as primary factors.

Although it is generally accepted that phonetic ability diminishes after puberty, a small number of individuals can reach native-like foreign language proficiency when they are adults (Selinker, 1972). Individuals with musical ability seem to be more likely to reach native-like foreign language proficiency (Christiner, 2020; Christiner et al., 2021, 2023, 2022b). In this study, we conducted regression analyses to examine potential relationships between auditory phonological pattern recognition—a key competence for predicting new vocabulary retention (Meara, 2005)—and musical skills. The results demonstrate that both musical variables are significantly associated with auditory phonological pattern recognition, supporting previous research suggesting a close connection between musical skills and language processing abilities (Besson et al., 2011; François et al., 2013; Kraus & Chandrasekaran, 2010; Thompson et al., 2004). In a previous study, we used a similar measures; however, in each trial the speech samples came from multiple speakers and spanned various languages, which introduced greater variability between language segments. Notably, this increased variability was not associated with musical ability (Christiner, 2020). This suggests that musicians may particularly benefit from perceiving subtle differences when the speech material is more homogeneous, as was the case in the present study. Consequently, it may indicate that musically skilled individuals possess a more finely tuned auditory system, enabling them to better discriminate and remember phonetic elements, consistent with other research (Patel, 2007). Previous studies have also revealed a significant link between musicians' performance and their ability to accurately assess their own musical skills (Inabinet et al., 2021), implying that individuals with an accurate perception of their musical abilities tend to excel in related tasks. Musically trained individuals often develop superior self-assessment skills because they regularly rely on auditory feedback—tuning their own performance against competing instruments—which hones their awareness of acoustic nuances and strengthens their ability to evaluate their musical abilities accurately. This self-estimation may contribute to auditory phonological pattern recognition, as it reflects the awareness of individuals regarding their auditory processing ability. These findings align with previous research indicating that musical ability is associated with foreign language production tasks, such as pronunciation and accent imitation (Christiner et al., 2021, 2023, 2022a).

The regression analysis also revealed that STM capacity was a significant predictor of performance on the auditory phonological pattern recognition measure. This was expected, as STM capacity is associated with a number of different language learning processes, such as speech production, reading, comprehension, and vocabulary acquisition (Coumel et al., 2023; Gathercole et al., 1997; Gathercole, 2006; Gathercole & Baddeley, 1990, 1993; Korecky-Kröll et al., 2019). Evidence supporting the role of phonological STM stems from clinical investigations, where patients experiencing impairments in the phonological loop exhibit difficulties in acquiring foreign languages (Baddeley, 2010; Baddeley et al., 1998). Conversely, individuals who excel in learning second languages, such as individuals with extensive foreign language capacity, demonstrate a sophisticated phonological STM, which stands as a key predictor contributing to their proficiency in multiple languages (Christiner et al., 2021; Papagno et al., 1991; Papagno & Vallar, 1995).

In previous research, we found that the education of the parents was associated with individual differences in respect to the acquisition of language aptitude measures (Christiner, 2020). In this study we wanted to replicate earlier findings. In order to generate a comprehensive SES variable, we used a weighted composite score of the educational status of the primary and the secondary caregiver during the participants' childhood (see Section 2). The regression analysis demonstrated that the SES variable was a significant predictor of auditory phonological pattern recognition in this study. The findings of this study show that the SES into which individuals are born appears to influence individual differences in language ability. We believe this has two crucial implications. First, it shows that the effect of socioeconomic disparities on language skills may persist well into adulthood (Hoff, 2013) even for abilities which should be less affected by external factors. Therefore, efforts should be directed towards the development of concepts or programs designed to mitigate socioeconomic differences. Second, aptitude measures need to be rethought. From a developmental perspective, it is said that aptitude is more likely to manifest itself at a younger age, as well as aptitude is transformed into talent through practice (Gagné, 2008). This view aligns with notions on aptitude and talent proposed by other researchers that emphasize the importance of practice and environmental factors in developing exceptional abilities and simultaneously challenges the notion of innate talent (Howe et al., 1998). SES research suggests that differences in the quality of interactions and opportunities for language learning influence individual differences in language performance (Cartmill et al., 2013; Rowe, 2018). Pedagogically, this underscores the importance of equitable access to enriched language experiences and targeted interventions early in life to support and enhance auditory and phonological abilities, helping to close SES-related gaps in language ability. However, when considering that aptitude develops into talent through practice (e.g., through quality interactions), the distinction between language aptitude and talent remains ambiguous, as aptitude measures may inadvertently capture the influence of prior environmental factors, such as socioeconomic status.

We performed various group analyses to examine the relationships between musical, language, and other variables, with a particular emphasis on auditory phonological pattern recognition (Q3). The participants were also divided into four groups, musically naïve participants with and without extensive foreign language experience. Extensive foreign language training meant that they spoke and understood at least two foreign languages on a high level. The findings of our study have shown that, as expected, participants with extensive foreign language experience outperformed the group without extensive training. Notably, this superiority was observed within the subset of participants who did not have musical training (see Table 3 and Supplemental Table S3). It could be suggested that extensive foreign language training has a significant effect on auditory phonological pattern recognition. This finding also corresponds to previously published studies (Gresele et al., 2013; Kaushanskaya & Marian, 2009) and is further illuminated when considered from a developmental linguistics perspective. Individuals possess the capacity to differentiate among all languages at birth, yet this ability diminishes gradually during the initial year of life in favor of their maternal mother tongue (Eimas et al., 1971). Therefore, researchers have suggested that the only capacity ultimately subject to a critical period is phonetic ability, which diminishes after puberty (Moyer, 2014). It is generally accepted that multilinguals are accustomed to perceiving and distinguishing between different phonemes, which can be beneficial when encountering new languages (Hell & Mahn, 1997; Kaushanskaya & Marian, 2009; Papagno & Vallar, 1995). Individuals who speak more foreign languages are exposed to a broader range of acoustic features across languages. This exposure can make them more familiar with various sound patterns and accents, allowing them to adapt more quickly to novel phonetic structures (Spinu et al., 2023; Tremblay & Sabourin, 2012). Therefore, the learning of foreign languages may have a similar effect on perceptual skills as we see with musical training. However, when comparing the auditory phonological pattern recognition performance of all four groups in this study, the two groups with musical training outperformed the two non-musical groups. Therefore, it could be suggested that musical training has a greater impact on recognizing phonological patterns in unintelligible auditory stimuli than extensive foreign language training.

Studies have shown that both musicians and non-musicians show immediate improvements in frequency discrimination after brief musical training. This suggests that perceptual skills may be highly responsive to practice (Kishon-Rabin et al., 2001). Indeed, research on the impact of musical training on the neural encoding of speech has shown that a two-year structured program of music training is enough to enhance how the brain encodes speech (Tierney et al., 2013). However, our auditory phonological pattern recognition task parallels musical aptitude assessments in two key ways. First, both use novel stimuli—whether specific melodies or unfamiliar phonological sequences—that participants have not encountered before, requiring them to rely solely on acoustic processing. Second, both require similar cognitive processes—such as sound classification, pattern recognition, and memory—whether for short musical pieces or multi-syllabic speech units. The strong association between auditory phonological pattern recognition and musical capacity suggests that incorporating music training into foreign language education may not only enhance auditory processing skills in linguistic contexts but also significantly support and accelerate learning processes.

While auditory phonological pattern recognition was notably better among musicians, effectively distinguishing them from non-musicians, STM capacity did not significantly contribute to this differentiation, as it loaded onto the discriminant function slightly too low.

Limitations

Although the majority of participants were native German speakers, and the remaining had diverse L1 backgrounds (Latvian, Croatian, English, Russian, Bosnian, Bulgarian, Polish), some native languages may share structural or phonological features with the target languages (Mandarin, Japanese, Farsi, Tagalog), potentially facilitating recognition performance. Future studies should therefore control for L1 background to account for cross-linguistic transfer effects.

5. Conclusions

In conclusion, our research has shown that STM capacity, SES, foreign language proficiency, and musical training have an influence on auditory phonological pattern recognition. However, the findings reveal that in initial foreign language learning settings, musical ability is a better predictor of auditory phonological pattern recognition than extensive foreign language training. Our findings also have implications for foreign language learning. Since musical training influences the ability to recognize auditory phonological pattern in unintelligible stimuli, it shows that fine-grained auditory perceptual skills should be trained in initial foreign language learning settings. In addition, the results of our study also have implications for educational policies and interventions, as they highlight the importance of addressing socio-economic disparities in access to quality language education. Educational programs should promote efforts to offer additional support and resources to individuals from lower SES backgrounds. Furthermore, our study also suggests that reevaluating aptitude measures is essential for comprehending the intricate dynamics among individual abilities, environmental influences, and language proficiency.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/languages10110272/s1. Section S1 includes the questionnaires used to classify participants as musicians or non-musicians. Section S2 provides the questions used to classify participants as either having or not having extensive foreign language training. Section S3 provides information on how socioeconomic status (SES) was determined, as well as details on participants' educational background, parental education, and the main caretaker during childhood. Section S4 and Table S1 present detailed information on the auditory phonological pattern recognition measure, including example items in Mandarin. Section S5 contains the list of abbreviations (Table S2), the Gabriel-corrected post-hoc analysis of a separate ANOVA for the auditory phonological pattern recognition measure (Table S3), and descriptive statistics for the variables under consideration (Table S4).

Author Contributions: Conceptualization, M.C.; methodology, M.C. and C.G.; formal analysis, M.C.; investigation M.C. and C.G.; resources, M.C.; writing—original draft preparation, C.G. and M.C.; writing—review and editing, C.G. and M.C.; visualization, M.C. and C.G. All authors have read and agreed to the published version of the manuscript.

Funding: Open access funding by the University of Graz.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board of the University of Graz GZ. 39/106/63 ex 2024/25.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data used in this study are available upon request from the corresponding author due to privacy and due to ethical reasons.

Acknowledgments: M.C. is a recipient of an APART-GSK Fellowship of the Austrian Academy of Sciences at the Department of Psychology at the University of Graz. The authors also acknowledge the financial support of the University of Graz. We sincerely thank all participants who generously contributed their time and effort to this study. We also express our deep gratitude to the anonymous reviewers, whose insightful and constructive comments helped us to further improve and advance this work.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Abrahamsson, N., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning*, *59*(2), 249–306. [CrossRef]

Alexander, J. A., Wong, P. C. M., & Bradlow, A. R. (2005, September 4–8). *Lexical tone perception in musicians and non-musicians*. Interspeech 2005 (pp. 397–400), Lisbon, Portugal.

Archibald, L. M. D. (2017). Working memory and language learning: A review. *Child Language Teaching and Therapy*, *33*(1), 5–17. [CrossRef]

Baddeley, A. D. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, *36*(3), 189–208. [CrossRef] [PubMed]

Baddeley, A. D. (2010). Working memory. *Current Biology*, *20*(4), R136–R140. [CrossRef] [PubMed]

Baddeley, A. D., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*(1), 158–173. [CrossRef]

Baran-Łucarz, M. (2016). The relationship between language anxiety and the actual and perceived levels of foreign language pronunciation. *Studies in Second Language Learning and Teaching*, *1*(4), 491. [CrossRef]

Besson, M., Chobert, J., & Marie, C. (2011). Transfer of training between music and speech: Common processing, attention, and memory. *Frontiers in Psychology*, *2*, 94. [CrossRef]

Calvo, A., & Bialystok, E. (2014). Independent effects of bilingualism and socioeconomic status on language ability and executive functioning. *Cognition*, *130*(3), 278–288. [CrossRef]

Calvo, M. G., & Eysenck, M. W. (1996). Phonological working memory and reading in test anxiety. *Memory (Hove, England)*, *4*(3), 289–305. [CrossRef]

Cartmill, E. A., Armstrong, B. F., Gleitman, L. R., Goldin-Meadow, S., Medina, T. N., & Trueswell, J. C. (2013). Quality of early parent input predicts child vocabulary 3 years later. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(28), 11278–11283. [CrossRef]

Christiner, M. (2018). Let the music speak: Examining the relationship between music and language aptitude in pre-school children. In S. M. Reiterer (Ed.), *Exploring language aptitude: Views from psychology, the language sciences, and cognitive neuroscience* (Vol. 16, pp. 149–166). English Language Education 16. Springer Nature.

Christiner, M. (2020). *Musicality and second language acquisition: Singing and phonetic language aptitude phonetic language aptitude* [Dissertation thesis, University of Vienna].

Christiner, M., Bernhofs, V., & Groß, C. (2022a). Individual differences in singing behavior during childhood predicts language performance during adulthood. *Languages*, *7*(2), 72. [CrossRef]

Christiner, M., Bernhofs, V., Sommer-Lolei, S., & Groß, C. (2023). What makes a foreign language intelligible? An examination of the impact of musical ability and individual differences on language perception and how intelligible foreign languages appear. *Journal of Intelligence*, *11*(3), 43. [CrossRef]

Christiner, M., Gross, C., Seither-Preisler, A., & Schneider, P. (2021). The melody of speech: What the melodic perception of speech reveals about language performance and musical abilities. *Languages*, *6*(3), 132. [CrossRef]

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

Christiner, M., & Reiterer, S. M. (2018). Early influence of musical abilities and working memory on speech imitation abilities: Study with pre-school children. *Brain Sciences*, *8*(9), 169. [CrossRef]

Christiner, M., Renner, J., Christine, G., Seither-Preisler, A., Jan, B., & Schneider, P. (2022b). Singing mandarin? What elementary short-term memory capacity, basic auditory skills, musical and singing abilities reveal about learning mandarin. *Frontiers in Psychology*, *13*, 895063. [CrossRef]

Christiner, M., Serrallach, B. L., Benner, J., Bernhofs, V., Schneider, P., Renner, J., Sommer-Lolei, S., & Groß, C. (2022c). Examining individual differences in singing, musical and tone language ability in adolescents and young adults with dyslexia. *Brain Sciences*, *12*(6), 744. [CrossRef]

Coumel, M., Groß, C., Sommer-Lolei, S., & Christiner, M. (2023). The contribution of music abilities and phonetic aptitude to L2 accent faking ability. *Languages*, *8*(1), 68. [CrossRef]

DeKeyser, R. M., Alfi-Shabtay, I., & Ravid, D. (2010). Cross-linguistic evidence for the nature of age effects in second language acquisition. *Applied Psycholinguistics*, *31*(3), 413–438. [CrossRef]

Delogu, F., & Zheng, Y. (2020). Beneficial effects of musicality on the development of productive phonology skills in second language acquisition. *Frontiers in Neuroscience*, *14*, 618. [CrossRef]

Dewaele, J.-M., & Wei, L. (2012). Multilingualism, empathy and multicompetence. *International Journal of Multilingualism*, *9*(4), 352–366. [CrossRef]

Doughty, C., Campbell, S. G., Mislevy, M. A., Bunting, M. F., & Bowles, A. R. (2008). Predicting near-native ability: The factor structure and reliability of the Hi-LAB. In T. Prior, Y. Watanabe, & S.-K. Lee (Eds.), *Selected proceedings of the 2008 second language research forum* (pp. 10–31). Cascadilla Proceedings Project.

Dörnyei, Z. (2006). Themes in SLA research. *AILA Review*, *19*, 42–68. [CrossRef]

Dörnyei, Z., MacIntyre, P. D., & Henry, A. (Eds.). (2015). *Motivational dynamics in language learning*. Second Language Acquisition 81. Multilingual Matters.

Eccles, R., van der Linde, J., Le Roux, M., Swanepoel, D. W., MacCutcheon, D., & Ljung, R. (2021). The effect of music education approaches on phonological awareness and early literacy: A systematic review. *The Australian Journal of Language and Literacy*, *44*(1), 46–60. [CrossRef]

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, *171*(3968), 303–306. [CrossRef]

Franco, F., Suttora, C., Spinelli, M., Kozar, I., & Fasolo, M. (2021). Singing to infants matters: Early singing interactions affect musical preferences and facilitate vocabulary building. *Journal of Child Language*, *49*(3), 552–577. [CrossRef]

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

Gagné, F. (2004). Transforming gifts into talents: The DMGT as a developmental theory. *High Ability Studies*, *15*(2), 119–147. [CrossRef]

Gagné, F. (2008). Building gifts into talents: Brief overview of the DMGT 2.0. *Gifted*, *152*, 5–9.

Gardner, R. (1990). Attitudes, motivation, and personality as predictors of success in foreign language learning. In T. Parry, & C. W. Stansfield (Eds.), *Language aptitude reconsidered* (pp. 179–221). Prentice Hall.

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, *27*(4), 513–543. [CrossRef]

Gathercole, S. E., & Baddeley, A. D. (1990). The role of phonological memory in vocabulary acquisition: A study of young children learning new names. *British Journal of Psychology*, *81*(4), 439–454. [CrossRef]

Gathercole, S. E., & Baddeley, A. D. (1993). *Working memory and language. Essays in cognitive psychology*. Erlbaum.

Gathercole, S. E., Hitch, G. J., Service, E., & Martin, A. J. (1997). Phonological short-term memory and new word learning in children. *Developmental Psychology*, *33*(6), 966–979. [CrossRef]

Gordon, E. (1989). *Advanced measures of music audiation*. GIA.

Gresele, A. D. P., Garcia, M. V., Torres, E. M. O., Santos, S. N. D., & Costa, M. J. (2013). Bilingualism and auditory processing abilities: Performance of adults in dichotic listening tests. *CoDAS*, *25*(6), 506–512. [CrossRef]

Groß, C., Bernhofs, V., Möhler, E., & Christiner, M. (2023). Misjudgement of one's own performance? Exploring Attention Deficit (Hyperactivity) Disorder (ADHD) and individual difference in complex music and foreign language perception. *International Journal of Environmental Research and Public Health*, *20*(19), 6841. [CrossRef] [PubMed]

Groß, C., Serrallach, B. L., Möhler, E., Pousson, J. E., Schneider, P., Christiner, M., & Bernhofs, V. (2022). Musical performance in adolescents with ADHD, ADD and Dyslexia—Behavioral and neurophysiological aspects. *Brain Sciences*, *12*(2), 127. [CrossRef] [PubMed]

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate data analysis* (7. Auflage, Pearson new international ed.). Pearson Custom Library; Pearson.

Han, Z. (2013). Forty years later: Updating the fossilization hypothesis. *Language Teaching*, *46*(2), 133–171. [CrossRef]

Hauser, R. M., & Warren, J. R. (1997). 4. Socioeconomic indexes for occupations: A review, update, and critique. *Sociological Methodology*, *27*(1), 177–298. [CrossRef]

Hell, J. G. V., & Mahn, A. C. (1997). Keyword mnemonics versus rote rehearsal: Learning concrete and abstract foreign words by experienced and inexperienced learners. *Language Learning*, *47*(3), 507–546. [CrossRef]

Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: Implications for closing achievement gaps. *Developmental Psychology*, *49*(1), 4–14. [CrossRef]

Howe, M. J., Davidson, J. W., & Sloboda, J. A. (1998). Innate talents: Reality or myth? *Behavioral and Brain Sciences*, *21*(3), 399–407; discussion 407–442. [CrossRef]

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, *61*(4), 343–365. [CrossRef]

Inabinet, D., de La Cruz, J., Cha, J., Ng, K., & Musacchia, G. (2021). Diotic and dichotic mechanisms of discrimination threshold in musicians and non-musicians. *Brain Sciences*, *11*(12), 1592. [CrossRef]

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*(1), 60–99. [CrossRef]

Jørgensen, H. (2008). What is a talent, how is it identified, and how is it developed? In *Precollege Education and Talent Development*. European Association of Conservatoires (AEC).

Kaushanskaya, M., & Marian, V. (2009). The bilingual advantage in novel word learning. *Psychonomic Bulletin & Review*, *16*(4), 705–710. [CrossRef]

Kishon-Rabin, L., Amir, O., Vexler, Y., & Zaltz, Y. (2001). Pitch discrimination: Are professional musicians better than non-musicians? *Journal of Basic and Clinical Physiology and Pharmacology*, *12*(2), 125–144. [CrossRef]

Kofler, M. J., Singh, L. J., Soto, E. F., Chan, E. S. M., Miller, C. E., Harmon, S. L., & Spiegel, J. A. (2020). Working memory and short-term memory deficits in ADHD: A bifactor modeling approach. *Neuropsychology*, *34*(6), 686–698. [CrossRef]

Korecky-Kröll, K., Dobek, N., Blaschitz, V., Sommer-Lolei, S., Boniecki, M., Uzunkaya-Sharma, K., & Dressler, W. U. (2019). Vocabulary as a central link between phonological working memory and narrative competence: Evidence from monolingual and bilingual four-year-olds from different socioeconomic backgrounds. *Language and Speech*, *62*(3), 546–569. [CrossRef] [PubMed]

Kragness, H. E., Swaminathan, S., Cirelli, L. K., & Schellenberg, E. G. (2021). Individual differences in musical ability are stable over time in childhood. *Developmental Science*, *24*, e13081. [CrossRef] [PubMed]

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. Nature reviews. *Neuroscience*, *11*(8), 599–605. [CrossRef]

Lenneberg, E. H. (1967). The biological foundations of language. *Hospital Practice*, *2*(12), 59–67. [CrossRef]

Li, S., Xu, Q., & Xia, R. (2019). Relationship between SES and academic achievement of junior high school students in China: The mediating effect of self-concept. *Frontiers in Psychology*, *10*, 2513. [CrossRef]

Linck, J. A., Hughes, M. M., Campbell, S. G., Silbert, N. H., Tare, M., Jackson, S. R., Smith, B. K., Bunting, M. F., & Doughty, C. J. (2013). Hi-LAB: A new measure of aptitude for high-level language proficiency. *Language Learning*, *63*(3), 530–566. [CrossRef]

Liu, Z. (2022). The interplay of English as a foreign language learners' interest, self-efficacy, and involvement. *Frontiers in Psychology*, *13*, 837286. [CrossRef] [PubMed]

Loosli, S. V., Buschkuehl, M., Perrig, W. J., & Jaeggi, S. M. (2012). Working memory training improves reading processes in typically developing children. *Child Neuropsychology*, *18*(1), 62–78. [CrossRef] [PubMed]

Ludke, K. M. (2018). Singing and arts activities in support of foreign language learning: An exploratory study. *Innovation in Language Learning and Teaching*, *12*(4), 371–386. [CrossRef]

Ludke, K. M., Ferreira, F., & Overy, K. (2014). Singing can facilitate foreign language learning. *Memory & Cognition*, *42*(1), 41–52. [CrossRef]

Meara, P. (2005). *LLAMA language aptitude tests: The manual*. Lognostics.

Milovanov, R. (2009, August 12–16). *Musical aptitude and foreign language learning skills: Neural and behavioural evidence about their connections*. ESCOM 2009: 7th Triennial Conference of European Society for the Cognitive Sciences of Music, Jyväskylä, Finland.

Milovanov, R., & Tervaniemi, M. (2011). The interplay between musical and linguistic aptitudes: A review. *Frontiers in Psychology*, *2*, 321. [CrossRef]

Moyer, A. (2014). What's age got to do with it? Accounting for individual factors in second language accent. *Studies in Second Language Learning and Teaching*, *3*, 443–464. [CrossRef]

O'Malley, M. J., & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. Cambridge University Press. [CrossRef]

Onwuegbuzie, A. J., Bailey, P., & Daley, C. E. (1999). Factors associated with foreign language anxiety. *Applied Psycholinguistics*, *20*(2), 217–239. [CrossRef]

Papagno, C., Valentine, T., & Baddeley, A. D. (1991). Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory and Language*, *30*(3), 331–347. [CrossRef]

Papagno, C., & Vallar, G. (1995). Verbal short-term memory and vocabulary learning in polyglots. *Quarterly Journal of Experimental Psychology Section A*, *48*(1), 98–107. [CrossRef]

Partanen, E., Kivimäki, R., Huotilainen, M., Ylinen, S., & Tervaniemi, M. (2022). Musical perceptual skills, but not neural auditory processing, are associated with better reading ability in childhood. *Neuropsychologia*, *169*, 108189. [CrossRef] [PubMed]

Pastuszek-Lipinska, B. (2008, June 29–July 4). *Influence of music education on second language acquisition*. Proceedings of Acoustics (pp. 5125–5130), Paris, France.

Patel, A. D. (2007). *Music, language, and the brain*. Oxford University Press.

Peretz, I., Vuvan, D., Lagrois, M.-É., & Armony, J. L. (2015). Neural overlap in processing music and speech. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *370*(1664), 20140090. [CrossRef] [PubMed]

Robinson, P. (2019). Effects of individual differences in intelligence, aptitude and working memory on adult incidental SLA. In P. Botana (Ed.), *Doing SLA research with implications for the classroom* (pp. 211–266). DeKeyser.

Rowe, M. L. (2018). Understanding socioeconomic differences in parents' speech to children. *Child Development Perspectives*, *12*(2), 122–127. [CrossRef]

Scharine, A., Cave, K., & Letowski, T. (2009). Auditory perception and cognitive performance. In *Helmet-mounted displays: Sensation, perception and cognition issues*. U.S. Army Aeromedical Research Laboratory. [CrossRef]

Seither-Preisler, A., Parncutt, R., & Schneider, P. (2014). Size and synchronization of auditory cortex promotes musical, literacy, and attentional skills in children. *Journal of Neuroscience*, *34*(33), 10937–10949. [CrossRef]

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, *10*, 1–4. [CrossRef]

Singleton, D., & Lengyel, Z. (Eds.). (1995). *The age factor in second language acquisition*. Multilingual Matters.

Skehan, P. (2012). Language aptitude. In S. Gass, & A. Mackay (Eds.), *The Routledge handbook of second language acquisition* (pp. 381–395). Routledge Handbooks in Applied Linguistics. Routledge.

Skehan, P. (2019). Theorising and updating aptitude. In P. Botana (Ed.), *Doing SLA research with implications for the classroom* (pp. 69–93). DeKeyser.

Spinu, L., Hwang, J., & Vasilita, M. (2023). Differences between monolinguals and bilinguals in phonetic and phonological learning and the connection with auditory sensory memory. *Brain Sciences*, *13*(3), 488. [CrossRef]

Swaminathan, S., & Schellenberg, E. G. (2020). Musical ability, music training, and language ability in childhood. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(12), 2340–2348. [CrossRef]

Swaminathan, S., Schellenberg, E. G., & Venkatesan, K. (2018). Explaining the association between music training and reading in adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(6), 992–999. [CrossRef]

Szabó, F., Polonyi, T., & Abari, K. (2019). Foreign language learning and low socio-economic status. *PedActa*, *9*(2), 37–46. [CrossRef]

Tewes, U. (1994). *Hamburg-Wechsler-intelligenztest für erwachsene-revision 1991 (HAWIE-R)*. Huber.

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2004). Decoding speech prosody: Do music lessons help? *Emotion*, *4*(1), 46–64. [CrossRef]

Tierney, A., Krizman, J., Skoe, E., Johnston, K., & Kraus, N. (2013). High school music classes enhance the neural processing of speech. *Frontiers in Psychology*, *4*, 855. [CrossRef]

Tremblay, M.-C., & Sabourin, L. (2012). Comparing behavioral discrimination and learning abilities in monolinguals, bilinguals and multilinguals. *Journal of the Acoustical Society of America*, *132*(5), 3465–3474. [CrossRef]

Wen, Z. (2012). Foreign language aptitude. *ELT Journal*, *66*(2), 233–235. [CrossRef]

Wen, Z., Biedroń, A., & Skehan, P. (2017). Foreign language aptitude theory: Yesterday, today and tomorrow. *Language Teaching*, *50*(1), 1–31. [CrossRef]

Wen, Z., & Skehan, P. (2011). A new perspective on foreign language aptitude research: Building and supporting a case for "working memory as language aptitude". *Ilha do Desterro*, *60*, 15–44. [CrossRef]

Werker, J. F., & Tees, R. C. (2005). Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Developmental Psychobiology*, *46*(3), 233–251. [CrossRef] [PubMed]

Wrembel, M. (2015). Metaphonological awareness in multilinguals: A case of L3 Polish. *Language Awareness*, *24*, 60–83. [CrossRef]

Yang, Y., Chen, X., & Xiao, Q. (2022). Cross-linguistic similarity in L2 speech learning: Evidence from the acquisition of Russian stop contrasts by Mandarin speakers. *Second Language Research*, *38*(1), 3–29. [CrossRef]