Correlations between musical and Japanese phonetic aptitudes by native speakers of English

Hitomi Nakata

School of Linguistics & Applied Language Studies, The University of Reading

Abstract. Music and language both utilise the auditory and oral channels to perceive and produce patterns of sounds in the time domain, and the possibility of similarities in processing has attracted the attention of researchers. Various approaches have been taken in the investigation of this interdisciplinary area, however, few studies have observed the potentially shared domains in a specific language and music. By analysing aptitude in a new (second) language and in a new piece of music, it may be possible to claim more 'feature-specific' correlation between music and language. This was the starting point of this empirical study. This study examines the relationship between aptitude in perceiving and producing rhythm (particularly syncopation), appropriate time duration, and pitch in music and aptitude in corresponding features of a second language: gemination, vowel lengthening, and pitch accentuation in Japanese, respectively. The results show some correlations between rhythm (in music) and geminates (in phonetics) obtained from 20 adult native speakers of English, which led to further investigation into the cognitive aspects of perceiving, integrating and interpreting sound input.

1. Introduction

This study investigates the correlation between music aptitude and second language (L2) aptitude for adults, based on empirical data obtained from adult native speakers of English. The experiment investigated the hypothesis that three phonetic features in Japanese as a new language may conceivably be correlated with three properties in music. Below is the list of the possibly correlated subcategories investigated in the test:

(1)	Music		Phonetics in Japanese
	Rhythm (syncopation)	↔	Geminates (double consonants)
	Time	↔	Lengthened (double) vowels
	Pitch	↔	Pitch accentuation

The test was given to 20 native speakers of English who had had no formal training in music or Japanese. Both the music and phonetic tests were designed to measure the subjects' abilities to discriminate subtle

Reading Working Papers in Linguistics 6. 1-23.

differences in minimal pairs. First, subjects' sensitivity to the geminates in Japanese and syncopation in music was measured. These two subcategories involve a unit of time which is felt but which has no audible onset. Second, the ability to detect suspended notes in music and lengthened vowels in Japanese words was tested. Finally the ability to discriminate between the slight differences in pitch both in music and in Japanese was measured.

The issue of the relationship between music and language has been debated since the beginning of the twentieth century. Enthusiasm for attempting to quantify human abilities (such as IQ) seems to have declined after 1980, however. Nonetheless, some scholars have attempted to demonstrate possible correlations between music and various languages. The conditions of such experiments and approaches vary. Overall, the main focus has been on phonological or phonetic features in language, while in music, both rhythm and pitch have been examined, as these are considered to be two central constituents of the musical domain. The details of the previous studies are discussed in Section 2 including their pros and cons which bear consequences for the hypotheses developed in this study.

The present study is organised as follows: In the Section 2, the discussion is developed with regard to the possibilities of shared properties between music and language. Previous studies are examined critically in the light of the presumed hypothesis of this study. The research method and materials used in the experiment are presented in Section 3. The reasons behind the design of my subtests in music and phonetics are explained. The section describes the approaches used in constructing test-items, and gives details about procedures, including subject-information. Finally, the results of the experiment are presented and discussed in Section 4.

2. Similarities and differences between language and music

2.1 General issues of the relationship between music and language

Both music and language are realisable through the vocal channel and both are perceived aurally as sound vibration. They are both expressive, showing feelings and meaning. However, there has been considerable work to show their distinctiveness and independence from each other, pointing out that the relationship, if it exists, is only marginal, and the process and mechanisms by which they work are dissociated (Gardner, 1983). Carroll (1981) also claims that there is not sufficient evidence to prove a strong relationship between language abilities and musical abilities, noting, for example, the studies of Arellano & Draper (1972) or Leutenegger, Mueller & Wershow (1965) which, he claims, fail to demonstrate a language-music correspondence. This section explores the possible relationship between language and music, reexamining the evidence demonstrated in the previous studies.

If information processing is operated through humans' auditory function for both music and language, it seems possible that there are at least some components which music and language share in terms of the process of gathering input and interpreting information. Dowling & Harwood (1986) suggest a similarity between abilities in experienced musicians and listeners who are unaware of the intellectual processes involved in making/hearing music and native speakers' judgements about their own languages, i.e. native speakers' implicit competence. Both music lovers and native speakers of a language are rarely able to describe the grammaticality of language or the rules governing language or music. All this processing is at a totally unconscious level and is related to cognitive aspects of perceiving, integrating and interpreting the meaning of what is heard.

Everyone listens to music either attentively or unconsciously, just as almost all humans speak their native language. In this sense, it can be said that both music and languages are universal and basic phenomena, shared by the vast majority of people in the world. In addition to the notion of universality, both music and language can be *Culture* specific, in the sense that a listener is unavoidably influenced by what s/he hears or is exposed to. It should be noted that the term *Culture* used here is not with a small 'c': culture; but with a capital 'C'; Culture, representing artefacts including music that is widely accepted and appreciated, beyond the boundaries of region and time.

Thus, it does not mean that all Russian people will enjoy Tchaikovsky's opera to the same degree, just because they were born in Russia. Rather, it means that opera may be appreciated by those who have been exposed to the values and the environment which includes the 'opera-culture', no matter where they are. The seven-note tonal scale developed in Europe has prevailed in many (even non-Western) countries, enabling people all over the world to become familiar with so-called western classical and pop music from childhood. As a result, it is quite possible to measure a certain type of music ability, if at least an elementary-level music education in the Western tonal system has been given. Also it is obviously possible to measure certain music abilities without considering the specific language used by the subjects. In other words, one may possibly possess a certain musical competence without being influenced by the first language (hereafter L1)'s musical qualities such as intonation, intensity, or pitch.

Before going on to a detailed description of more evidence for and against a relationship between music and language, I will review the discussion of cognitive and psychological properties which may be distributed similarly in the realms of language and music.

One might suppose that music is something to feel, sense and appreciate by means of universal human emotion which transcends cultures. But when we think of unfamiliar, non-Western classical music such as Balinese gamelan or Indian traditional music performed with the sitar, we have difficulties identifying the stream of sounds. (Dowling & Harwood, 1986). To be able to appreciate the meaning of a piece of the music, we must know at least the basic structure of the music, if it is not familiar. In other words, perception and interpretation of music requires some mental activities that can often be based on a 'once-conscious-but-became-unconscious' operation. For example, when we listen to a piano concerto by Chopin, we unconsciously perceive the structure of the music such as time signature, expected melody lines performed by the piano along with the matching orchestration, the change of tempo or volume, and utilisation of a variety of motifs. All these factors included in a piece of music become familiar from childhood, and hence become unconscious knowledge providing the ability to access to a new but similarly structured piece of music.

This could be a good analogy of learning an L2, particularly for adults who start learning L2(s) after establishing their L1 structure. Having developed the 'familiar' speech/music structure, when learners/listeners listen to the L1 or to familiar music, they can anticipate what is coming next, comprehend the meaning and the context simultaneously, and even make further associations using memories beyond the context of the given input. These mental activities cannot occur when listening to unfamiliar structures of language or music: a listener cannot simultaneously perceive, interpret and contextualize unfamiliar auditory input. In terms of processing information, language and music can be said to share the same dimensions and aspects in cognitive and psychological organisation.

Another analogy can be drawn between methods of assessing abilities in music and language. Almost all conventional test batteries seem to employ a segmental or 'bottom-up' approach for measuring aptitude in music including Seashore (1919), Wing (1948), Bentley (1966), and the Modern Language Aptitude Test by Carroll & Sapon (1959). But do these tell us anything about the subjects' abilities to appreciate and understand the meaning of a phrase of music or language in larger chunks?

This issue is raised by Gardner (1983), in which he points out that '… from the basic phonological level, through a sensitivity to word order and word meaning, to the ability to appreciate larger entities like stories, … it is possible to examine sensitivity to individual tones or phrases, but also to look at how these fit together into larger musical structures which exhibit their own rules of organisation.' (ibid.:108). Here he suggests that the method of assessing capacity should be re-examined, taking account that broader and suprasegmental approaches are needed for analysis. That is, more attention should be given to 'top-down' prosodic approaches, rather than focusing on individual segments of music and language.

Such studies would certainly require more careful and complicated consideration regarding the design of the material, and a more sophisticated system of evaluating the outcome, than merely measuring the segmental properties of language and music. So far there have been few studies measuring abilities in this large context. The next two sections will look at the few existing studies and examine the evidence they provide for and against a correlation between music and language.

2.2 Evidence in support of a relationship between music and language

In the development of the theory of equivalencies of music and language, Jackendoff & Lerdahl (1980) present the fundamental concept of a generative approach to tonal music (Generative Music Theory), deriving from generative linguistics. One of the important elements in their framework is an attempt to explore the relationship between the prosodic structure in language and the notation structure in music. They suggest that the organisation of stress in larger units of language (e.g. feet) presented by Liberman & Prince (1977) is associated with the underlying structure in music; namely, the series of consonance designated from the relatively less important ornamental (pitch-) events to the more essential parts in order to form the main structure of the passage. This perception of dynamism in both musical and linguistic structures is noteworthy with regard to the whole prosodic organisation including the prediction of stress placement. Thus, the theory developed by Jackendoff & Lerdahl deals with abstract structures rather than surface features.

Tumanov (1986) suggests that there are possible correlations between categories of language and vocal music. He accepts the theory of Jackendoff & Lerdahl, and attempts to apply their phrasal structure to a more specific area of music: songs. The discussion turns to the effort of the composers in adjusting the musical development to the meaningful phrase (the verbal phrase). He concludes that language used in vocal music shows an equivalent development in music. His study is based on only Russian traditional songs, however, the manifestation of the concurrent motion of music and words may exemplify the potential similarities in both properties.

Arellano & Draper (1972) suggest a possible relationship between music and language. Their study of ten-year-olds investigates aptitude in music (measured with the Seashore Test) and the results of Spanish learning over a six-week period of audio-lingual instruction. Spanish as L2 was assessed by a test consisting of two major parts: speech production and comprehension. All subcategories were further divided into multiple-choice Q/As in Spanish for comprehension; and phone, intonation, stress, etc. for production. They concluded that in child L2 acquisition, there is a possibility of interaction of music and L2 learning abilities, and mutual reinforcement may occur.

The study also demonstrates interesting evidence of children's abilities in both areas. Mental activities in children may work differently from information processing in adults. Significant insight could thus be gained by examining the two groups (children and adults) and acquisition of other languages. Yet, there have been few studies comparing language-specific features of a particular language to sub-categories of music.

2.3 Evidence against a relationship between music and language

According to Gardner (1983), parallels exist in music and language intelligence in '… the abilities of individuals to discern meaning and importance in sets of pitches rhythmically arranged and also to produce such metrically arranged pitch sequences as a means of communicating with other individuals.' (ibid.:98). Gardner questions, however, whether the once-shared expressive medium continued or developed over the history of evolution. He argues that music intelligence is a separate intellectual competence, as is language intelligence. He claims that music has its own cognitive representation. He further discusses how some researchers (musicologists) have '… attempted to apply Chomsky's analysis of the generative structure of language to the generative aspects of musical perception and production.' (ibid.:125), although he doubts that parallels can be drawn between language and music, given the largely regular grammaticality of language as compared with the great irregularities allowable within music.

Gardner (1983:110-111) discusses Bamberger's (1982) theory of musical competence, whereby '… musical thinking involves its own rules and constraints, and cannot simply be assimilated to linguistic or logical-mathematics thinking', insisting that the musical domain is not interchangeable with that of language, and that children's approach to music is, 'intuitive, based solely on what is heard irrespective of any theoretical knowledge of music.'

Gardner (ibid.:117) introduces another interesting conclusion offered by Deutsch (1975) that '… the mechanisms by which pitch is apprehended and stored are different from the mechanisms that process other sounds, particularly those of language.' Deutsch draws this conclusion on the basis of a study showing that non-melodic verbal material did not interfere with melodic material (2% errors), whereas other interposed tones drastically influenced memory (40% errors). This might be evidence for the existence of a music-particular faculty which does not coexist with other properties such as language.

Gardner (ibid.:125) characterises this independent faculty as 'autonomous intellectual competence', claiming that '… not all aspects of language are directly analogous to music: for example, the whole semantic aspect of language is radically underdeveloped in music, and the notion of strict rules of 'grammaticality' are once again extraneous in music, where violations are often prized.' He emphasises that semantic aspects are partially independent, so it is possible that the other area(s) of language could be comparable to some equivalent domain of music, to some extent, possibly in phonetic and phonological areas which presumably correspond to each other.

2.4 Hypothesis

The preceding discussion suggests that correspondences may be found between musical and phonetic features, i.e. if a subject performs well in the perception or production of rhythm, s/he will be good at perceiving or producing geminates as well. Likewise, good performance may be seen for lengthened vowels if the subject scores highly on note duration in music, and similarly for pitch accent in Japanese and tonal memory.

The first possible similarity is thus, between syncopation and Japanese gemination. Syncopation involves an unrealised beat resulting from tying one note to the next rather than beginning a new note. An example in musical notation follows:

(2)


This is sometimes referred to as placing the ‘Auftakt’ on the offbeat, and the result is an apparent ‘limping’ or suspended rhythm. We assume that the composer and the listener know the rhythm of the music being performed and are aware that there is a beat on which there is no new note, so that there is a continuity of perceived rhythm, even though the actual rhythm is interrupted.

Geminate consonants in Japanese, have the same effect: one might say that there are two consonants in a row, but the first is not released, so the first runs into the second without a new initiation. The second consonant retains the duration it would have had in the first place, thus producing a ‘limping’ pattern. An example from Japanese follows:

(3) [so.k.ku.ri] ‘to look exactly alike’

In this sense of ‘unproduced but deserving attention’ syncopation and long consonants are thus similar. This leads to our first hypothesis, namely that good performance on the production and perception of syncopation will be highly correlated with the production and perception of long sounds.

Long vowels (which can be thought of as double vowels, much like consonantal geminates) can also introduce unrealised onsets in Japanese. For example, in what is written here as two vowels in the Japanese word for ‘coffee’, *koohii,* each of the ‘o’s has a one-vowel duration, but the onset of the second vowel is not realised; hence a long vowel. This is because Japanese has a mora-based syllable structure where the following rules apply:

(4) a. A long consonant (geminate) is counted as two moraic units.
b. A long vowel (or diphthong) is counted as two moraic units.
c. A moraic nasal ‘n’ can be independently counted.

When learners of Japanese encounter this phenomenon, they must recognise that there are different time durations for lengthened vowels and single vowels. Therefore, it might be suggested that the ability of discriminating suspended sounds in music might be associated with the ability to perceive lengthened and non-lengthened vowels in Japanese words. Next we will see the third possible similarities between musical pitch and Japanese pitch accentuation.

Pitch is a fundamental factor in both music and Japanese. The accent patterns of Japanese are generally known as ‘pitch patterns’, which specify a high or low accent on each mora, similar to the high-low movement of notes in music. The examples below illustrate how high (H) and low (L) pitches are combined in words that have the same numbers of mora.

(5) a. ha si ‘chopsticks’ b. ha si ‘bridge’
H L L H
c. se n se i ‘a teacher’ d. se n se i ‘an oath’
L H H L L H H H

This representation of high/low pitch pattern differs from that of strong and weak stress pattern in English, and learners of Japanese often confuse this accentuation difference. Hypothesis 3 is thus that listeners who can discriminate the (slight) difference between pitches in music also have a high degree of sensitivity to pitch accentuation in language, and vice versa.

The next two sections explore the procedures and results of the experiment conducted on the possible relations between three domains of music (rhythm, time duration, and pitch) and the conceivable correlates of these in Japanese phonology (geminate, lengthened vowels and pitch accentuation). Discussion and analysis of the experiment follow.

3. Methods and materials

3.1 Music material

In order to measure musical abilities corresponding to the phonetic abilities assessed here, three categories were selected: rhythm (including syncopation), time duration, and pitch discrimination. 10 pairs for each subcategory were constructed to measure subjects’ ability to discriminate differences.

For the selection of these categories, I referred to the description of tests in Shuter-Dyson (1981) Appendix. For constructing the test items, I consulted (among others) The Seashore Measures of Musical Talents (1919), Bentley Measures of Musical Ability (1966), and Wing Standardized Tests of Musical Intelligence (1948). Each part was further divided into two or three subdivisions. Throughout all the sections, each pair contained minimal contrasts in rhythm, time duration (only one note is longer or shorter in the second pattern than in the first.), and pitch (only one note has a different pitch in the second pattern than in the first). Example materials are described in detail below (See also the Appendix).

For the first rhythm section, I adopted the style of rhythm discrimination of both Seashore (1919) and Wing (1948), whereby testees are asked if paired rhythm patterns are the same or different (hereafter S/D) (Seashore, 1919). Wing's (1948) test includes another instruction asking about which is the better version, if different. I omitted this question about preference on the grounds of its subjectivity. Instead, I added a production element to the rhythm test, in which subjects were asked to reproduce either the first or the second pattern (as instructed) of the same 10 pairs.

For the time duration, I modified Seashore's test, in which two tones are played and the testee is asked if the second tone was longer or shorter than first. I produced a more complex version, whereby the stimulus for each pair consisted of one measure of 4-8 notes. The subjects were asked to answer three questions: (1) Are the two sequences equal length? (2) If different, which note in the sequence (of the second pattern) sounded different? (3) Was it shorter or longer than the first pattern? In analyzing the data, these (1)-(3) were integrated into one figure as mean of the perception of the time duration section.

Finally for the pitch imagery (tonal memory) test, 10 pairs of sequences of 3-8 notes were played. This differs slightly from Seashore, in which testees are instructed that they will hear 30 pairs of tonal sequence, 10 items each of three-, four-, and five- tones, and they are asked which note is different. In some pairs both sequences have the same tone. Subjects were required to judge whether or not the pitch of the second sequence was the same as or different from the first. If different, the subjects were asked to specify which note in the sequence (first, second, etc.) was different.

3.2 Phonetics (Japanese) material

The material for testing phonetic abilities consists of, again, three sections: geminates, lengthened vowel, and pitch accent. All these are language-specific features in Japanese phonetics. Each section was subcategorized into two parts: perception and production. The production part was coincidentally a similar procedure to Neufeld's (1987:324) experiment. His subjects had also had no contact with the target languages (Chinese, Japanese, and Eskimo). Throughout all the subtests, no meaning or grammatical clues were given to the subjects, in order to have them concentrate on the sounds. The following describes the characteristics of the three selected features from the Japanese phonetic system.

First the geminate was tested. In the test, one minimal pair given is *gakki* 'musical instrument' and *gaki* 'children' (slang), the latter having no geminate. The length of the utterances ranged from two to four syllables for both the perception and production elements of the test. Subjects were expected to decide whether or not each pair contained a difference and mark on the pre-lettered answer sheet: S (same) / D (different). Four of the ten pairs contained words that did not differ, in order to control for subjects potentially answering randomly.

Japanese vowel lengthening was the second category in the test. Again, a minimal pair was provided so as to measure whether subjects could discriminate between, e.g., *tooru* 'to go through' and *toru* 'to take'. The length of the words ranged from two to six syllables, and again four of the pairs consisted of two words that were the same.

The third specific phonological feature of Japanese in the test was pitch accent. Some words have the same sound structure but distinctive pitch (high or low) on each mora, which leads to distinctive meanings. For this section, subjects were asked to draw a contour marking the pitch rising and falling in the second pattern. An example of a contour was visually presented in order for them to understand how they should mark. Below is an example with the boxed words indicating the tape-recorded material heard by the subjects for all three categories.

				Gloss
(6)	1. Geminate	Perception	macci / maci	matches / town
		Production	kakko / kako	parenthesis / past
	2. Lengthened Vowel	Perception	moteru / mooteru	popular / motel
		Production	waku / waaku	frame / work
	3. Pitch accent	Perception	a si / a si	legs / reed
		Production	ki ru / ki ru	to cut / to wear

3.3 Procedure

All subjects were tested in a quiet room (either at their home or in a university classroom), either individually or as a group. It was possible for the perception tests to be given to a group, but the production part was recorded individually. First, subjects were given both the music and the phonetics perception tests. Individual production part came afterwards. While individual recordings of the production part were carried out, the waiting subjects filled out the questionnaire. The test time was 40 minutes for group perception and 15 minutes for the individual production part.

The musical test items had been recorded using a mini-keyboard. For the rhythm and pitch sections, the piano timbre was used, whereas the organ timbre was used for time duration section in order to have the length of each note maintained enough to be distinguishable. The time signature was not told to subjects. All three sections consisted of the same number of a variety of measures: 2/4 (×2), 3/4 (×2), 4/4 (×4) and 6/8 (×2); altogether 10 pairs. Care was taken to assign the same number of notes within pairs throughout all 30 pairs, so as to facilitate identification of different notes. For the pitch section, two Japanese traditional melodies were also employed (the *Miyako-bushi* scale and a traditional folk song scale, for #7 and #8 respectively), with the idea of providing subjects with some non-Western tonal scales which were unfamiliar, like the Japanese sounds.

The phonetic test material (paired and isolated Japanese words) had been recorded by the examiner. All three parts consisted of a ‘conventional auditory discrimination procedure where subjects were asked to distinguish between phonemes in minimal pair contexts’ (cf. Neufeld, 1987:325). Only section three included extra work for the subjects, in marking the pitch (high/low) on the answer-sheet. All the material was presented without providing grammatical clues or semantic information. Subjects listened to the paired words to discriminate sounds.

The stimulus utterances were given only once for each pair. The tape material included an instruction after each pair saying, ‘first one’ or ‘second one’. Subjects were requested to imitate only the one as instructed.

The test material also did not include any visual information (except in the pitch accent test) such as alphabetical symbols, so as to avoid orthographical influence. For example, native speakers of English would probably produce the sound [r] if they saw the letter ‘r’, instead of the Japanese [ɾ]. In other words, the test excluded orthographic information so as to tap the subjects’ genuine aptitude for listening to and reproducing the sounds of an unfamiliar language.

3.4 Questionnaire

In the questionnaire, the subjects’ background in both music and language was surveyed. In the language part, questions were asked about familiarity with languages other than English, and degree of proficiency. The subjects’ experience of Japanese was also questioned, as well as their impression of how it sounds when a native speaker of Japanese speaks

English in order to see if subjects detect the Japanese sound features in non-native speech in English. The music part started by asking whether the subjects were musical at all followed by questions about any musical training received. This question is crucial, since any formal training in music could influence subjects' performance on the test. The next questions asked about music preferences and about the frequency of listening to music in general. Finally space was given for comment.

3.5 Subjects

All the participants – 11 females and 9 males – were recruited from the general public and from the University of Durham. The age range was from 21 to 55 years old. Only adults were targeted because one of the purposes of this study is to find out if it is possible to acquire L2 phonology after the critical period. Measuring a phonetic ability in a totally new language can be a predictor for a successful learner. The conditions for participating in this study were that the subjects should be native speakers of English, they should have had no significant exposure to Japanese, and they should have had no (or little) formal training in music.

4. Results and discussion

4.1 Results

Subjects' responses on the language aptitude tests were tape-recorded and assessed by the examiner (experienced in playing the piano and a native speaker of Japanese). If an utterance among the Japanese phonetic results was ambiguous (i.e., on the border-line of being acceptable or intelligible), it was double-checked by another native speaker of Japanese. Mistakes irrelevant to the particular section were ignored and subjects got points if the target element was produced correctly, e.g., [dzɪssɪn] was accepted for [jɪssɪn] even though the onset consonant is incorrect, because the target element here was the geminate. Likewise, in the overall scoring, attention was paid to whether the target phenomena were recognized and performed correctly throughout the production parts of both the music and phonetics tests. In the perception part, however, these details could not be analyzed due to the nature of the multiple-choice answer format.

Table 1, below, shows the test scores. As mentioned in the previous section, the production test was conducted only for one subsection (rhythm) in the music part and for all three subsections in the phonetic part. All the subcategories of music and phonetics included 10 question items. The abbreviations stand for as follows: Pe – Perception, Pr – Production, S/D – Same or Different, N# – Note number, S/L – Short or Longer, PM – Pitch Marking, and MS – Mean of Sum.

Table 1 Scores for the two aptitude tests: music and phonetics

	Music							**Phonetics**						
	Rhythm		**Time-duration**			**Pitch**		**Geminates**		**Vowel-length.**		**Pitch Accent**		
	Pe	Pr	Pe			Pe		Pe	Pr	Pe	Pr	Pe		Pr
			S/Dc	N#	S/L	S/D	N#					S/D	PM	
MS(%)	**86**	**41**	**85**			**78**		**84**	**82**	**74**	**94**	**75**		**87**

The figures in Table 1 are the ones that were used for calculation of the means, standard deviation and correlation, to be discussed later. Some groups were further divided into two to three subcategories. First, the perception of pitch in music was divided into two sub-tasks: (1) whether the paired tone was judged to be the same or different, and (2) if different, which note in the sequence was the source of the difference. The S/D and N# scores were combined for calculation of the mean because it was considered that this combination would give a more precise indication of subjects' aptitude. The same idea of combining scores was applied in the time duration sections. Second, the pitch marking scores in the pitch accent perception were ignored for the calculation. This is because the intricate instructions for marking pitch contours onto the sequenced syllables on the answer sheet resulted in drastically lower scores, compared with the S/D discrimination task. Including these scores would have yielded a misleadingly low score for pitch perception. I assumed that the ability to draw pitch contour lines was not a correlate of pitch perception ability.

The percentages given in Table 1 do seem to indicate a correlation between musical and phonetic ability: comparable scores are obtained for perception of rhythm and geminates (86% vs. 84%), and perception of pitch in music and phonetic pitch accent (78% vs. 75%), and perception of time duration and the mean of perception and production of vowel-lengthening (85% vs. 84%). Thus, the evidence shows the hypotheses presented in Section 2 were generally confirmed.

Also from the above data (at least from the mean of sum), subjects achieved surprisingly high scores for both parts: music and phonetics. The performance was particularly high for time duration in music (85% for perception only), and all the subcategories in phonetics (geminates: 83%, lengthened vowels: 83%, and pitch accent: 81%, all figures being means of perception and production). Interestingly, the means of two subcategories of phonetics were better in production than in perception (lengthened vowel: 74% for perception and 94 % for production; pitch accent: 75% for perception and 87% for production), whereas the scores for production in rhythm was considerably lower than for perception. Finally, the discrepancy between perception and production on geminates was quite low (only ±0.2 points difference) compared with the other sections where both perception and production were given, such as rhythm (±4.5 points difference), vowel-lengthening (±2.0 points difference), and pitch accent (±1.2 points difference).

In the Table 2 below, the deviation was calculated to see how the scores in each subcategory varied.

Table 2 Musical and phonetic aptitude of 20 adult English native speakers: means and standard deviations

Music	Adult NS (n = 20)	
	M	SD
Rhythm (Pe)	8.60	1.00
Rhythm (Pr)	4.05	2.04
Time duration (Pe)	8.48	1.34
S/D		
N#		
S/L		
Pitch (Pe)	7.78	1.60
S/D		
N#		

Phonetics	Adult NS (n = 20)	
	M	SD
Geminates (Pe)	9.35	1.04
Geminates (Pr)	7.35	2.19
Vowel Length. (Pe)	9.90	0.31
Vowel Length. (Pr)	8.70	1.59
Pitch Accent (Pe)	7.40	1.02
S/D		
PM		
Pitch Accent (Pr)	8.90	1.48

As Table 2 shows, higher variability is seen for production both of rhythm in music and of geminates in phonetics (s = 2.04 and s= 2.19, respectively) than for any other subcategories. This suggests that a wide range of scores were obtained on these tests (0-7 points in rhythm in music and 4-10 points for geminates in phonetics). In other words, the performance on both production tasks varied according to the individuals.

The least variable value was obtained for perception of vowel lengthening in phonetics (s = 0.31). In fact, the highest score of all (mean = 9.4) was obtained for this subcategory, and 19 subjects attained high scores, ranging from 8 to 10 points. The rest of the subcategories show moderate deviations which range between 1.02 and 1.60. Thus, it seems that there are slight correlations between rhythm in music and geminates in phonetics, in that both show high rates of variability in production ability. Also, in a similar sense, it may be uniformly less difficult to obtain high scores in vowel lengthening, although this may not be related to the hypothesised equivalency, time duration in music (s = 1.34), which shows relatively high variance.

To look for further support for the three hypotheses of conceivable relationships between rhythm vs. geminates, time duration vs. vowel lengthening, and pitch vs. pitch accent, let us consider correlations between music abilities and phonetic abilities. Table 3 shows these correlations.

Table 3 Correlations between music and phonetics (subcategorized into perception and production)

Phonetic Aptitude	Music Aptitude					
	Rhythm (Pe)	Rhythm (Pr)	Time (Pe)		Pitch (Pe)	
Geminates (Pe)	-0.15					
Geminates (Pr)		0.37				
Vowel-L (Pe)			0.12			
Vowel-L (Pr)						
Pitch (Pe)					-0.16	
Pitch (Pr)						

From Table 3, it seems clear that there were no significant correlations between music and phonetics for categories of perception and production. In fact, there were even negative correlations between rhythm perception and geminate perception (-.15) and pitch perception in music and pitch perception in phonetics (-.16), suggesting that a high score in one area predicts a low score in the other. On the other hand, measures of production show some significant correlations such as between rhythm production and geminate production (.37). This discrepancy between abilities to perceive and produce the same property (rhythm and geminates) may suggest that the human operation of perceiving input data and conveying it to the channel of processing to produce the data is complex. That is, listening, comprehending, internalising then reproducing the ‘once-conceptualized-information’ are all separate operations. Particularly in measuring the abilities of adult subjects, one may need to consider their analytical approach to any unfamiliar auditory input. However, this perception-production connection may not have been fully captured by the present test, as production in music had only one section:

rhythm. If subjects had been given production tasks in all the categories, the data might have been different. Thus, it may be more valid if each subcategory is considered as a whole, rather than divided.

Therefore, the most direct relationships between the main categories for music and phonetics were calculated. Table 4 displays the correlations in general between each three subcategories of music and phonetics.

Table 4 Correlations between music and phonetics aptitude measures

Phonetics Aptitude	Music				
	Rhythm		Time D.		Pitch
Geminates	0.60				
Vowel-lengthening			0.17		
Pitch					0.09

From Table 4, the first thing noted is that rhythm and geminates show the highest correlation .60. The other two correlations are much lower: time and vowel-lengthening: .17; and pitch (music) and pitch accent: .09. It seems that the correlation is obtained only between rhythm and geminates.

4.2 Discussion

The aptitude tests for both music and phonetics were given to 20 adults native speakers' of English. The hypothesis was that three specific subcategories of music might be correlated with three specific subcategories of phonetics. The results showed that the correlation between rhythm (especially in production) and geminates (also in production) was significant enough to confirm the hypothesis concerning those subcategories. Both the perception and production parts of the rhythm and geminate test results were not significantly correlated with other subcategories, particularly in the rhythm and time duration sections. This evidence suggests that these abilities may be distinctive from each other and deserve independent attention. In addition, a slightly significant correlation (.60) was seen between the integrated rhythm test and the geminate test (Table 4). The figure was relatively high compared with the other two (.17 and .09).

Thus it can be claimed that rhythmic recognition is more directly connected to humans' operating system of information processing. This argument may be supported by the results of this study showing that all the adult subjects (who were not trained in music, except for a few) used strategies to analyse new materials and to fully exploit their stored knowledge in order to process the new sounds.

These strategies of dealing with unfamiliar information could have worked when they encountered new language features such as geminates in Japanese. They might have had either preconceived ideas about Japanese or ideas derived from the previous perception part of the test material, or they might even have searched for some similar phonemes such as [t] of *kit* which resembles a geminated sound since the fricative [t] has to wait (for a moment) before being pronounced due to the tongue shift from the preceding velar [k]. To investigate these strategies, I gave the subjects some questions about their experience in Japanese. In fact, 11 subjects gave the answer ‘staccato’ to the question of what kind of image they had of Japanese. Subjects also might have drawn on past experience of hearing English spoken by Japanese people. In the questionnaire, 12 answered ‘Japanese use different sounds’ to the question of ‘What is your impression of Japanese (language), if you have heard any’. This implies that the subjects were not completely unbiased with respect to the unfamiliar language.

Prediction of a high correlation between these rhythmic aspects both in unfamiliar language and in music was borne out. On the other hand, the other two hypothesised correlations, time and vowel-lengthening, and pitch discrimination in both areas, were not supported by the results. This is a little surprising, because some researchers suggest that at least musical pitch can be associated with a property of language. For example, Arellano and Draper (1972) discuss Dexter’s study (1934) of university and high school students who learn French as their L2. The results showed that the better the subjects performed in music – pitch discrimination – the better they performed on accent in French. Dexter (1934) also mentions that the correlation increases the younger the subject. This suggests that all age differences need to be considered: not only adult-child division but also the smaller age divisions, e.g. teenagers and people in their twenties, etc. The fact that the present study included subjects from 21 to 55 might have been a factor behind the generally low correlations.

Leutenegger, Muller & Wershow (1965) also conclude that only Tonal Memory (one of the Seashore’s subtests) emerged as significant in predicting L2 acquisition success, in their study of measuring L2 proficiencies compared with all the Seashore tests given to two groups of university students learning Spanish and French. In addition, they cite factors including the subjects’ academic intelligence, and motivational influences as having an effect on L2 proficiency. Leutenegger et al. focus particularly on the correlation of Tonal Memory with general intelligence, adding '… it would seem that Tonal Memory is the basis to recognition, identification, and remembrance "over something longer than a few seconds"' (ibid.:30). This may suggest that at least some properties (including Tonal Memory) are stored in abstract form.

My subjects' performance on vowel lengthening was extremely good; however, this did not correlate with the hypothesised similarity to the assumed music equivalent; time-duration (.12). Their good performance may be due to their unconscious familiarity with a similar phenomenon in English – long vowels. The test materials were segmental but the phonetic test included prolonged sounds with a word, making this test closer to real language than the equivalent music test was close to real music. In the latter, subjects had to judge the length of sounds from only one note in isolation ('A' on the scale). This was probably artificial, and hence more difficult for the subjects. Providing bigger chunks in a suprasegmental context might encourage better performance both in a new language and in music. A bigger chunk provides learners more clues on which to apply strategies for analysing the material. In other words, if they are given only segmental information, they need to depend on their intuition which might or might not have been developed. But if given a larger context, they would have more input and hence the possibility of applying a broader range of knowledge and processing strategies.

This could be one of the explanations for the relatively low correlations in this study. In the experimental conditions, the subjects had to be quite intuitive and speedy, and they could not make use of the schemata for problem solving.

One other possible factor which might have caused the relatively low correlations could be that although an attempt was made to exclude subjects with a background in music or exposure to Japanese, it turned out that some of them had a lot of experience in musical training, e.g. two university students had more than 10 years' experience of playing the piano; and one subject had stayed in Japan for a year. The two subjects who had had music training succeeded in answering the item #6 and #9 of the music rhythm section correctly, in contrast with the others.

The second reason for the low correlation may be the test format itself. Almost all subcategories consisted of 'discrimination among sound patterns' which is a conventional well-established method but also a highly criticised measure. This may have induced a certain formality, i.e. a psychological effect in the subjects, and this resulted in relatively high scores throughout the test.

5. Conclusion

The results and discussion in Section 4 demonstrated that the ability to acquire music and the ability to acquire a new language are partially correlated. Of the subcategories examined, it seems that the aptitude to discriminate the syncopated rhythmic patterns is correlated with sensitivity to geminates in (a new) language. By contrast, the other two-hypothesised correlations, namely time vs. lengthened vowel, and pitch both in music and language were not supported by the results.

The findings point to various intriguing issues concerning implications for methodology in pronunciation teaching, and the psychological basis of comprehension of phonology and music.

In this study, all the subcategories required the subjects to respond immediately after the stimuli both for perception and production teaching, not testing proficiency but intuitive responses in a short time (5 sec. or so). The test style was like a production test in phonetics. Subjects imitated either the first word or the second, thus, in fact testing their short-term memory abilities, where they did not have time to analyse, interpret the sounds or make associations with familiar sounds.

If aptitudes for both language and music are tested on the assumption that memory is the very key to contributing to adults' capacities, the relevant abilities should be tested not only from short-term memory but also from long-term memory. Results might have been different if subjects had been given stimuli after longer intervals. As mentioned briefly in Section 2, the organisation of stress in larger units of language (e.g. feet) is associated with the underlying structure in music.

This assumption can be related to the implications for effective adult L2 learning. Adults L2 learning entails not only an immediate, intuitive understanding but also meta-linguistic strategies such as knowing abstract rules and processing the input data. In fact, some interesting phenomena were observed in the production tests in the Japanese phonetics section of the experiment. Some production errors were not from the subjects' L1 (English) but obviously from the subjects' effort in assimilating the 'new' sounds which they just heard.

A further question which cannot be ignored, is that concerning the segmental approach of the experiment. That is, as Gardner (1983) points out concerning music tests, this style of testing individual segments of a block of music does not satisfy some musicians who favour using more artistic and musical entities in tests, like those which humans naturally encounter. This suggests that perception of larger, more realistic segments should be measured.

Finally, it might also be of interest to investigate so-called ‘prodigies’ of either language or music. In this case, it would be important to systematise the test material to ensure that it measures their mode of processing, as it often can be the case that these talented (and usually very young) children depend on their ‘intuitive’ understanding of language or music, rather than on conscious effort of exploiting the type of knowledge which is usually measured by aptitude tests.

References

Arellano, S. I. & Draper, J. E. (1972). Relations between musical aptitude and second-language learning. *Hispania* **55(1)**. 111-121.

Bamberger, J. (1982). Growing up prodigies. the mid-life crisis. *New Directions for Child Development* **17**. 61-78.

Bentley, A. (1966). *Measures of musical abilities*. London: Harrap Audio-visual.

Carroll, J. B. (1981). Twenty-five years of research on foreign language aptitude. In Diller, K.C. (ed.). *Individual differences & universals in language learning aptitude.* Massachusetts: Newbury House Publishers, Inc.

Carroll, J. B. & Sapon, S. M. (1959). *Modern Language Aptitude Test (MLAT).* New York: The Psychological Corporation.

Deutsche, D. (1975). The organization of short-term memory for a single acoustic attribute. In Deutsch, D. & Deutsch, J.A. (eds.) *Short-term memory*. New York: Academic Press.

Dexter, E. S. & Omwake, K. F. (1934). The relation between pitch discrimination and accent in modern languages. *Journal of Applied Psychology* **18**. 271.

Dowling, W.J. & Harwood, D. L. (1986). *Music cognition*. California: Academic Press, Inc.

Gardner, H. (1983). *Frames of mind. The theory of multiple intelligences*. London: Heinemann.

Jackendoff, R. & Lerdahl, F. (1980). *A deep parallel between music and language*. Bloomington: Indiana University Linguistics Club.

Leutenegger, R. R., Mueller, T. H. & Wershow, I.R. (1965). Auditory factors in foreign language acquisition. *Modern Language Journal* **49**. 22-30

Liberman, M. & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry* **8**. 249-336.

Neufeld, G. G. (1987). On the acquisition of prosodic and articulatory features in adult language learning. In Ioup, G. & Weinberger, S. H. (eds). *Interlanguage phonology. The acquisition of a second language sound system*. Cambridge and Massachusetts: Newbury House Publishers.

Seashore, C. E. (1919). *Psychology of musical talent.* Boston: Silver Burdett.

Shuter-Dyson, R. (1981). Musical ability. In Deutsch, D. (ed.) *The psychology of music*. New York: Academic Press.

Tumanov, A. (1986). Music in L2 pedagogy. A comparative analysis of music and language. *Russian Language Journal* **136-137**. 35-54

Wing, H. (1948). Tests of musical ability and appreciation. An investigation into the measurement, distribution, and development of musical capacity. *The British Journal of Psychology*. Monograph Supplements.

Appendix: Examples of Music & Phonetic Aptitude Test

Phonetics (Perception)

		Gloss
1. Geminates	(1) macci / maci	matches / town
	(2) syokku / shoku	shock / job
	(3) sassuru / sasuru	to guess / to stroke
2. Lengthened vowels	(1) tori / torii	bird / a torii
	(2) yoosei / yosei	a fairy / momentum
	(3) moteru / mooteru	popular / motel
3. Pitch accent	(1) mici / mici	way / unknown
	(2) sake / sake	salmon / sake (to drink)
	(3) se n se i / se n se i	oath / teacher