Resolving Conflicting Linguistic and Musical Cues in Metric & Beat-Strength Perception of Songs

by Jieun Oh

Submitted to the Symbolic Systems Program (B.S.)

Stanford University

In fulfillment of the Honors requirement

June 11, 2008

Thesis Advisor: Jonathan Berger, Professor in Music

_____

Second Reader: Lera Boroditsky, Professor in Psychology

_____

Abstract

This study explores the interplay of accentuation in language and in music by examining situations in which implicit accents in a melody conflict with accentual patterns in the lyrics. While previous studies describe resolution of such conflicts in the context of accentuating melodic and rhythmic patterns according to the musical cues that result in incorrect text accentuation, the opposite phenomenon—in which lyric accentuation maintains its integrity when set with strongly conflicting musical accentuation—has only been speculated and artificially generated. This study examines a unique instance in which the Korean translation of the internationally popular song "Happy Birthday" renders its strong implicit musical accentuation subordinate to the conflicting linguistic features, resulting in a perceptual alteration of the intended metric structure of the song. A preliminary experiment using hand-claps suggests that native Koreans may perceive the first beat of the song as a strong beat contrary to what is implied by the musical rhythm and instructed by the notated musical score. A follow-up experiment using a more precise finger-tapping method confirms this tendency, thereby offering empirical evidence to the speculated possibility of text overriding music.

TABLE OF CONTENTS

CHAPTER I

Introduction & Motivation

1.1 OVERVIEW

Music and language are special cognitive abilities that uniquely characterize humans. There are similarities between the two: they both have a logical syntactic structure, their primary goal is to communicate ideas and feelings, and they affect other areas of cognition. However, it is unclear whether these are merely some apparent high-level analogies (Baryshev, 1989), or whether music and language actually share much of the same processing, particularly during the early developmental stages of our brain (Trehub & Trainor, 1993; Fassbender, 1996; McMullen & Saffran, 2004).

McMullen & Saffran (2004) summarize developmental parallels between the faculties of language and music. However, neurological evidence suggests cortical separation of music and language in adults. For instance, language impairment is more often associated with injury of the left temporal lobe, while a music-specific processing deficit of amusia is generally related to damage to the right temporal lobe (Peretz & Coltheart, 2003). The apparent contradiction between the neurological data suggesting modularity and behavioral results suggesting parallels between the linguistic and musical systems is, to date, not reconciled.

Exploring how music and language interact in the context of associated music and lyrics may offer insights into the concurrent use of the two systems. This study thus explores the interplay of spoken language and music by seeking examples in which the accentual inference

derived from melodic or rhythmic factors in the musical setting is inconsistent with the accentual inflection[1] of the spoken words as set to music. While previous studies offer many examples in which the musical accentuation dominates over text accentuation when the two are in conflict, the possibility of the opposite phenomenon—in which lyric accentuation maintains its integrity when set with strongly conflicting musical accentuation—has only been speculated and artificially generated. This study examines a rare instance, found through the Korean version of the well-known song "Happy Birthday to You"[2], in which the musical features become subordinate to the conflicting linguistic features, resulting in a perceptual alteration of the intended metric structure of the song.

1.2 **PERCEPTION OF BEAT, METER, AND RHYTHM**

1.2.1 Definitions[3]

Beat is the basic pulse underlying mensural music. In a typical notated western music, each measure starts with a strong beat (downbeat) and ends with a weak beat (upbeat). Downbeats create a periodic occurrence and are usually given articulation through dynamic

[1]The term "accentuation" used in the linguistic context of this paper should be interpreted loosely to mean various types of emphasis (stress) given to certain syllables in a word, achieved through variation in pitch, loudness, and/or duration. Note that the ways stress manifests itself in speech are highly language dependent (Wikipedia entry on *Stress*: http://en.wikipedia.org/wiki/Lexical_stress)

[2] "Happy Brithday to You" is the most popular song in the English language according to the 1998 Guinness Book of World Records, and it is widely considered to be the most often-heard song universally, having been translated into at least 18 languages according to Robert Brauneis' *Copyright and the World's Most Popular Song*. The melody comes from the song "Good Morning To All" written by Patty Hill and Mildred J. Hill in 1893, and much attention has been drawn to the song's copyright status, as some argue that unauthorized public performances of the song are technically illegal unless royalties are paid to it. (Wikipedia entry on *Happy Birthday to You*: http://en.wikipedia.org/wiki/Happy_Birthday_to_You)

[3] Unless otherwise noted, this section is taken from Grove Music Online Dictionary: www.grovemusic.com

increase or lengthening of durational value. In contrast, upbeats (also known as *anacruses*) are weaker anticipatory notes preceding and precipitating the downbeat.

The grouping of strong and weak beats into larger units constitutes *Meter*. Meter is commonly categorized as duple or triple (according to whether the measure is organized in groups of two or three beats). Meter functions as a dynamic temporal framework for the production and comprehension of musical durations.

Rhythm is generically defined to be a 'movement marked by the regulated succession of strong or weak elements' (Oxford English Dictionary). Roughly speaking, it is the arrangement of notes according to their relative duration and relative accentuation, but the precise definition of rhythm in a musical context is tough to pin down.

1.2.2 Factors Contributing to the Perception of Musical Meter

What is the mechanism through which music performers and listeners come to infer the musical meter? Intuitively, perception and recognition of the musical meter would involve grouping of beats into repeating accentuation patterns. Yet the process through which we are able to recognize meter is often not so trivial: Hannon, Snyder, Eerola, & Krumhansl (2004) suggest that perception of the musical meter involves considerations to multiple melodic and temporal features in music. Furthermore, a study by Berger and Gang (1998) discusses the following contributors to contextual bias that influence meter cognition: biological or experiential preference for a given metric schema, the influence of metric cues heard prior to audition of the work, and veridical expectations based upon style or genre. In this manner, the perceptual task of determining the musical meter can be quite complicated and not so clear-cut.

The perception of musical meter involves perception of *downbeats*, which mark the beginning of the metric cycle. Wright (2007) mentions that "the downbeat, which is always a

metric accent, is not necessarily a *phenomenal accent*," where a metric accent is defined to be "time points in music that are perceived as accented by virtue of their position within a metrical scheme (Clarke 1999)," and a phenomenal accent is defined to be "points of local intensification caused by physical properties of the stimulus such as changes in intensity, simultaneous note density, register, timbre, or duration (Clarke 1999)." This distinction is crucial for our study, which examines cases in which downbeats are inferred based on the perception of phenomenal accents (including those derived from linguistic prosody), which may or may not be in alignment with metric accents suggested from the notated musical score. Finally, Wright remarks that "the perception of downbeat and metric accent in general can be completely subjective and is culturally learned," and the goal of this study is precisely to quantify the differences in the subjective downbeat and metric perception among people of varying linguistic, and inevitably cultural, background.

1.2.3 Subjective Meter: Preference of Duple over Triple

In studies involving beat, rhythm, and meter perception, it is important to take note of humans' general tendency to prefer duple over triple meters. Numerous experimental evidences have suggested an innate human preference for duple groupings, including a study by Bolton (1894), which first reported a prevailing tendency in almost all of his fifty subjects to group by four, given an isochronous sequence of equidistant tones—even though they could be perceived as consisting of groups of 2, 3, 4, or none.

In addition, a recent study by Bergeson & Trehub (2005) explored 9-month-old infants' perception of auditory temporal sequences, and found that infants detected change in pattern with

context of duple, but not triple, meter. In this manner, studies have consistently shown humans' preference for the duple meter.

1.3 PERCEPTUAL AMBIGUITIES

1.3.1 Ambiguous Meter

There are instances in music, referred to as "polymeter," in which phenomenal accents generate competing metric interpretations of a musical surface, resulting in an ambiguous metric perception that continue for a relatively prolonged period (Berger & Berger, 2004). The term "monophonic polymeter" is used to describe a situation in which the perceived polymeter is created by conflicting cues within a single melodic line. JS Bach's *Sarabande, Cello Suite #3* serves as an example of monophonic polymeter because listeners of this music are often conflicted with multiple possibilities in determining the piece's metric structure, and the interpretation would depend on the extent to which emphasis is placed on the first as opposed to the second beat of the phrase.

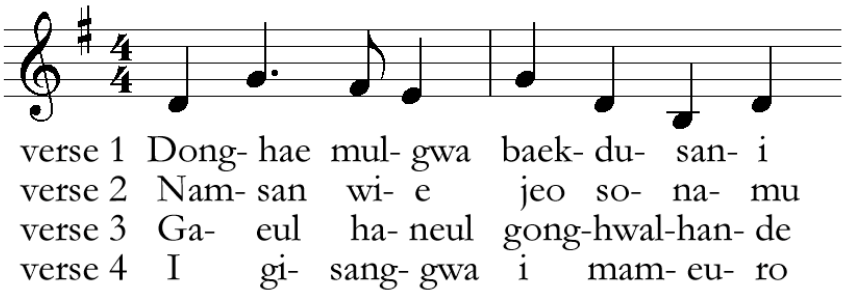
1.3.2 Ambiguous Downbeat

Directly related to the phenomenon of monophonic polymeter are situations in which conflicting accentuation cues make it difficult for the listener to locate the downbeat. Finding the downbeat is less problematic for the performer reading music from a notated score because the downbeat, by definition, is the first beat of each measure (which is clearly notated through the use of a bar line in music). However, when a performer intentionally de-emphasizes a downbeat

or accentuates an upbeat, the listener becomes surprised, and possibly even confused, by the unexpected change in accentuation patterns.

A common example of inducing ambiguous downbeat is through the use of a hemiola. Hemiola is a metrical pattern in which two bars in simple triple time (3/2 or 3/4 for example) are articulated as if they were three bars in simple duple time (2/2 or 2/4)[4], an altered accentuation pattern that temporarily gives a sensation of ambiguity in the downbeat.

While most occurrences of ambiguous downbeat are intentionally generated as a way of violating the listeners' musical expectations, there are unintentional instances in which the listener cannot easily locate the downbeat. For example, this ambiguity is likely to arise in the first measure of the Korean National Anthem as shown:



<FIGURE 1: From the Korean National Anthem>

In this example, the intervallic and durational accents (and, in some instances, text) all work together to suggest that the music begins with an anacrusis. First, the interval of a rising fourth occurring over the first two beats emphasizes the second beat over the first. Second, the elongated rhythm on the second beat gives rise to a durational accent. Finally the lyrics, particularly in verse 4, is semantically grouped as a single-syllable word (corresponding to the beat 1) followed by a three-syllable word (corresponding to beat 2, 3, and 4) with a higher

[4] Wikipedia article on *hemiola*. http://en.wikipedia.org/wiki/Hemiola. This entry also notes that musicians commonly extend the definition of "hemiola" to include any occasion of a "three-against-two" metrical feel --- including some mixed meters and polyrhythms --- contrary to the word's original meaning

intonation assigned to the syllable in beat 2 when spoken; semantics and intonation of the text thus result in a natural accentuation of the second beat in the music, assigning a weaker beat to the first. All of these factors will contribute to a sensation of having an anacrusis at the beginning of the song.

1.4 INTEGRATING MUSIC AND LANGUAGE

1.4.1 Music and Language: Developmental Comparison

Many studies have drawn parallels between the cognitive processing of music and language in the human brain, especially during the infant developmental stages. Such similarities—at least on the surface—include the sound, prosodic, and grammatical structures of the two systems (McMullen & Saffran, 2004). Some argue that the two domains are "indistinguishable during the early stages, and [it is] only until later stages that they become more diversified" (Chen-Hafteck, 1997).

It is important to note, however, that this "indistinguishability" between the music and the language domain in early developmental stages may be an unfortunate artifact of the limitations in experimental methods for studying infants. Papousek and Papousek (1981) acknowledged the difficulty in distinguishing between singing and speech in early vocalization, as they found it difficult to distinguish between the development of intonation contour in speech and that of melodic contour in singing before the acquisition of first lexical items and one-word sentences. Dowling (1988) makes a similar remark by contending that the distinction between singing and speech cannot be made when the child is very young.

Regardless, notable studies suggest the significance of the relationship between music and language in children's singing. For instance, the singing of Venda, Ghanaian, Japanese, and Hungarian children have been found to reflect the characteristics of their languages (Blacking, 1967; Addo, 1996; Minami & Umezawa, 1990; Kalmar & Balasko, 1987). Furthermore, Chen-Hafteck (1996) showed that tonal language affects pitch accuracy in singing by comparing Cantonese-speaking children and English-speaking children.

1.4.2 Linguistic Prosody Influencing Musical Structures

Patel and Daniele (2003) provides an empirical basis to the claim that prosody of a culture's spoken language leaves an imprint on the instrumental music of a culture by quantitatively comparing English and French with the classical instrumental music of England and France. They were able to demonstrate that English and French musical themes are significantly different in their quantitative measure of rhythm, which also differentiates the rhythm of spoken English and French. This important study offers empirical evidence for the power of language to potentially shape music.

1.5 CONFLICTING CUES: LANGUAGE VS. MUSIC

Section 1.4 noted instances in which language and music interact in a more-or-less constructive manner. But because spoken language has its own accentuation and intonation patterns, text-setting lyrics to a piece of music in a manner that is consistent with both the musical and the linguistic features of sound can be tricky, creating room for perceptual ambiguities. Interestingly, when such ambiguity arises, we generally see that the text remains

subordinate to the tune; rarely do we find instances in which linguistic prosody dominate over the musical rhythm in the resulting output of singing. This study, however, offers a concrete example of the latter scenario.

1.5.1 Musical Features Dominating Over Linguistic Features

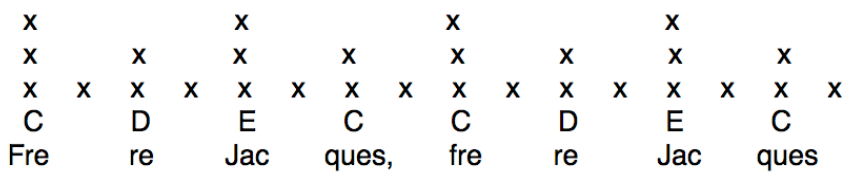
When conflicting inferential cues between tune and text are present, a singer generally follows the features in the tune. That is, when there is a mismatch between implied accentuation from the music and the lyrics, performers often follow the musical cues at the expense of accurate linguistic production.

An example of musical pitch overriding linguistic intonation is illustrated in a study by Mang (2007), which examines the speech-song interface of Chinese speakers. Mang finds that "mismatched linguistic tones were widespread among adults' (pop music) and children's songs (nursery rhymes)." Even though pitch is considered the most important carrier of information in Chinese, the study shows that Chinese speakers could choose to dispose of the lexical function when the musical context is the focus of conscious attention.

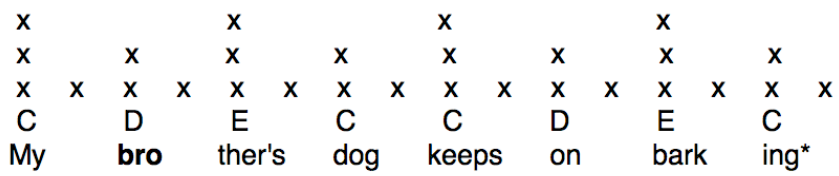
Another study by Ross (2003) demonstrates instances in Estonian songs in which the musical rhythm dominates over the prosodic shape of lyrics despite that "prosodic features (mostly duration) are used in Estonian in order to convey lexical and grammatical differences to listeners." Ross thus concludes that in singing, the prosodic structure of a language can in some cases be strongly subordinated to the musical structure, even though the result is unacceptable from the linguistic point of view.

1.5.2 Linguistic Features Dominating Over Musical Features

In this manner, studies have shown how linguistic features can be dominated by the musical structure in moments of conflict. In contrast, the opposite phenomenon—of the music being distorted for the sake of preserving the language—has only been suggested through an artificially construction: Morgan and Janda (1989) crafts a hypothetical scenario in which a text in English is dubbed over the original lyrics of the song "Frere Jacques" in such a way that the stress patterns of the musical line and the accent patterns in the English text are mismatched. **FIGURE 2** illustrate the tune with original lyrics (aligned), and **FIGURE 3** illustrates the tune with a made-up English text (misaligned):



<FIGURE 2: First phrase of the song "Frère Jacques" with original lyrics>



<FIGURE 3: Assigning English lyrics in a mismatched way that maintains the melody>

Morgan and Janda found that when subjects were forced to integrate the mismatched English lyrics to the existing tune, most subjects did not follow the pattern in **FIGURE 3**, but rather distorted the melody—often by inserting an anacrusis prior to the first note of the existing song to shift the first four events leftward (**FIGURE 4**)—in order to avoid the mismatched positions and resolve the conflict.

```
        x               x               x               x
        x       x       x       x       x       x       x       x
x       x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
<G>     C       D       E       C       C       D       E       C
My      bro     ther's  dog     -       keeps   on      bark    ing
```

<FIGURE 4: Common resolution of the mismatch: inserting an anacrusis>

This study by Morgan and Janda demonstrates that it may be possible to have situations in which the music becomes subordinate to text when a mismatch arises between the two. Unfortunately, however, the example remains as artificial, and observations from a naturally-occurring example of text overriding music have not been made.

1.5.3 The Case for "Happy Birthday to You"

A serendipitous observation at a birthday party led to the realization that the Korean lyrics to the well-known song "Happy Birthday" has a different stress patterns than the original English text, resulting in a mismatch in accentuation patterns between the text and the tune. To more carefully explore this phenomenon, a preliminary experiment was conducted in an attempt to track the metric perception of the song for native English and native Korean speakers. The results of this study was taken one step further in a follow-up experiment that more precisely tracks beat-strength perception of the song, to compare the effects of translated text on the perceived metric structure of the song.

Chapter II

Preliminary Observation (2006) :

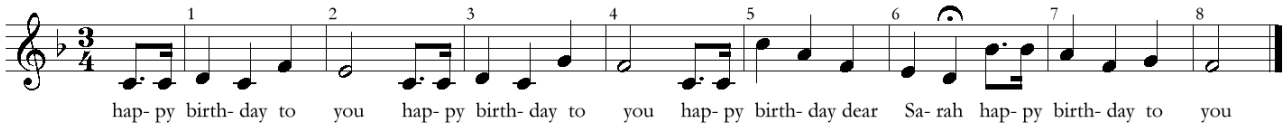
Downbeat Tracking through Clapping

2.1 INITIAL QUESTION & HYPOTHESIS

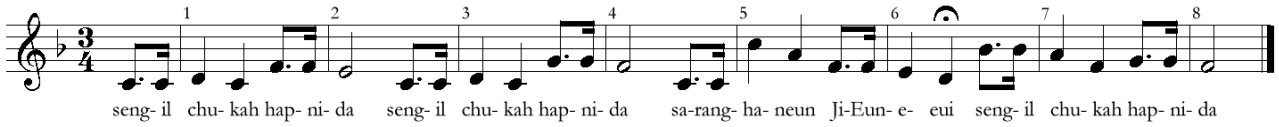
The goal of this preliminary study was to answer the question, "Can linguistic features of lyrics change the perceived metrical structure of a song?" This question is directly related to the larger issue of resolving conflicting inferential metric cues derived from features of the language and the tune—which together make up a song.

We observed singers' metric perception of "Happy Birthday" with respect to the language choice for the lyrics. We predicted that the Korean translated version of "Happy Birthday" would be perceived as starting on a downbeat, while the original English version, consistent with the notated music score, would be perceived as starting on an upbeat (anacrusis) as a result of the differences in the accentuation patterns in the English and Korean text.

The **FIGURE 5** and **FIGURE 6** show the original English version and the Korean translation version of the song "Happy Birthday" with its lyrics. Notice that the Korean version has two syllables on beat 3 of measure 1, 3, 5 and 7.



<FIGURE 5: Happy Birthday to You (in English)>



<FIGURE 6: Happy Birthday to You (in Korean)>

We now present a rough comparison of perceived beat-strength of each syllable of the song's lyrics in English and Korean, to motivate our predicted model of downbeat perception for the Korean translated version.

[Accentuation Patterns[5] for the English Lyrics]

Phrase 1: HAP-py **BIRTH-**day to **YOU**

Phrase 2: HAP-py **BIRTH-**day to **YOU**

Phrase 3: HAP-py **BIRTH-**day dear **SA-**RAH

Phrase 4: HAP-py **BIRTH-**day to **YOU**

[Accentuation Patterns[6] for the Korean Lyrics]

Phrase 1: **SENG-**il CHU-kah **HAP-**ni-da

Phrase 2: **SENG-**il CHU-kah **HAP-**ni-da

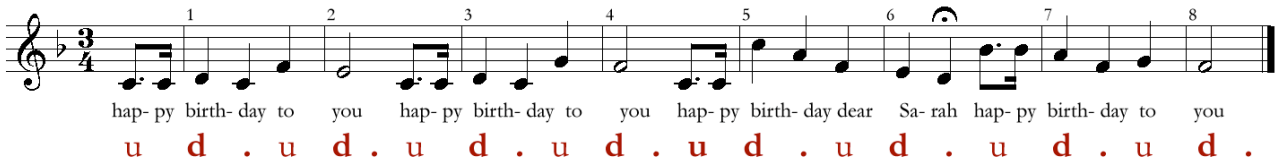
Phrase 3: **SA-**rang HA-neun **JI-**EUN-e-eui

Phrase 4: **SENG-**il CHU-kah **HAP-**ni-da

A correctly identified downbeat pattern according to the notated score is shown by **FIGURE 7**, where "d" stands for a downbeat and "u" stands for an upbeat:

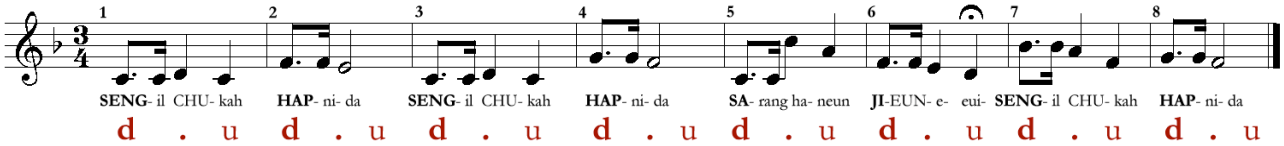
[5] Lower-case letters signify weak, capital letters signify strong, and bolded capital letters signify strongest.

[6] In contrast to English, Korean phonology is not based on stress accent; it is primarily intonation (pitch) based. The derived accentuation patterns reflect high-pitch assignment of the first syllable "seng" and "chu" of the two-syllable unit "seng-il" and the five-syllable unit "chu-kah-hap-ni-da."



<FIGURE 7: Downbeat-Upbeat Pattern in the Original Score>

According to the linguistic patterns shown above, we now predict that Koreans singing the song in Korean may *perceive* the metric structure of the song as the following:



<FIGURE 8: Predicted *perceived* metric structure for the Korean version>

Under this predicted model (**FIGURE 8**) that eliminates the anacrusis, the syllables that receive emphases in the Korean translation version would be very nicely aligned with the downbeats. In other words, we predicted that there would be a *shifting* of the downbeats in the Korean version of the song to accommodate the text.

2.2 METHOD: CLAPPING

Seven Korean bilinguals (about equally fluent in both Korean and English) and five native speakers of English (whose main language was English and could not speak any Korean) were asked to sing "Happy Birthday" and to simultaneously clap in places when they "felt the strong beat (downbeat)." Their singing and clapping were videotaped for further analysis.

2.3 RESULTS

2.3.1 Brief Summary of Findings

Excluding three subjects (2 Korean-English bilinguals and 1 native English speaker) who correctly perceived the song as having a 3/4 time signature that starts with an anacrusis, the general trend was that Koreans almost always clapped on the very first beat of the song, regardless of the language they sang in. What was not anticipated, however, was the enormous preference of clapping in a duple over triple pattern among Koreans (but not among native English speakers) in a way that ironically made their clapping pattern be inconsistent across the four phrases of the song. In other words, Koreans more-or-less clapped on every other beat, seemingly ignoring much of the melodic or syllabic cues present in the song. In contrast, native English speakers tended to rely on the lyrics and the melodic line (the cues for which are consistent with one another and therefore do not cause any conflict) to determine when to clap.

2.3.2 Observation of Native Korean Speakers (Korean-English Bilinguals)[7]

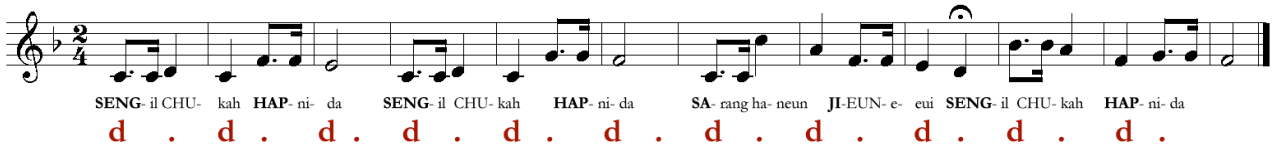
Surprisingly, Koreans in general clapped on every other beat, suggesting that they may have perceived the meter as either 2/4 or 4/4. Given this duple pattern of clapping, we were unable to confirm our original conjecture that Koreans would fail to perceive the anacrusis in the context of a triple meter.

Interestingly, this duple-patterned clapping persisted even when the Korean subjects sang in English. Two potential explanations to this phenomenon are offered here: (1) Not enough break period between the Korean and English trials prevented the subjects from switching out of

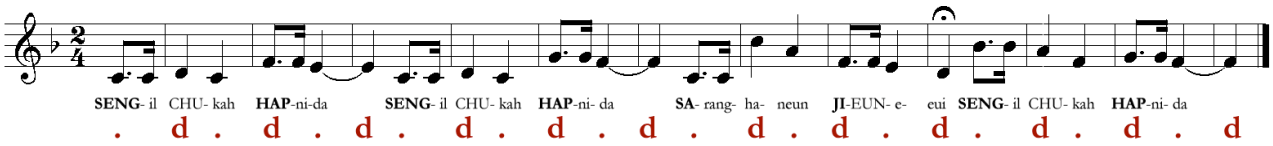
[7] Excluding two subjects who clapped "correctly" on the downbeat according to the notated score

the prime effects of the Korean version[8], and (2) The duple pattern of clapping is more of an automatic cultural habit[9], rather than a conscious effort of determining the downbeat.

Second unexpected finding was that Koreans added in an additional beat (or even two beats) at the end of each phrase in such a way that made their duple clapping highly inconsistent *across* the four phrases. Had they not included these additional beats, their clapping actually would have made the beat patterns for each of the phrases to look identical, achieving relative stability as illustrated by the hypothetical models in **FIGURE 9** and **FIGURE 10**. Despite that both of these alternatives allow for consistency across the four phrases, it would nonetheless feel quite unnatural to follow them because the stressed syllables in the text are rarely aligned with the downbeats.



<FIGURE 9: Hypothetical 2/4 model #1 of Korean version>

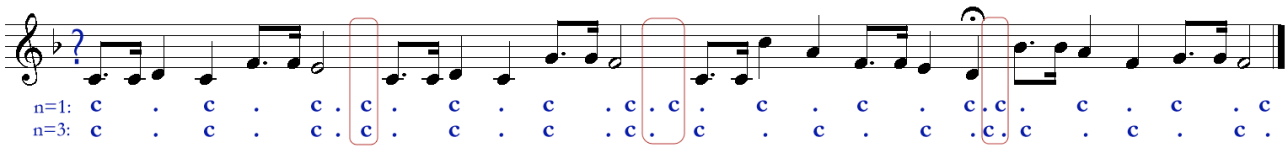


<FIGURE 10: Hypothetical 2/4 model #2 of Korean version>

[8] It would be easy to test this hypothesis by splitting Korean subjects into two groups so that one group starts with the Korean lyrics, and the other group starts with the English lyrics. Informal observation suggests that the duple patterned clapping occurs regardless, casting doubt on this hypothesis.

[9] We note that there is a great tendency in a Korean culture to clap along while singing songs and chanting cheers as a group, and it appears as though the clap patterns in these situations are almost always in duple meter

Interestingly, by inserting additional beat(s) at the end of each phrases—seemingly arbitrarily and inconsistently—phrase 2 (and occasionally phrase 3 and phrase 4, depending on the subject) actually ended up *recovering* the anacrusis that had been eliminated in the very beginning of the song. **FIGURE 11** shows the clapping patterns performed by four Korean subjects. A "c" denotes a clap and a "." denotes a beat without a clap. Insertions of extraneous beats at ends of phrases are shown inside red boxes.



<FIGURE 11: Clapping Results for Korean Subjects>

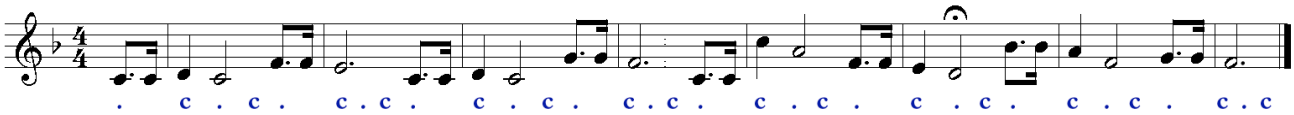
Both of these findings are quite puzzling. First, What made all Koreans to clap every other beat? We should be very cautious to conclude from the clap data that Koreans *perceived* "Happy Birthday" as a duple meter. That is, as suggested above, the duple-patterned clapping could simply have been an automatic cultural habit, a behavior that would be displayed in other circumstances involving singing, and a phenomenon *not* specific to "Happy Birthday." In this possible situation, the clap data would fail to provide meaningful insights into the subjects' metric perception of "Happy Birthday."

Second, why were the extra beat(s) added to the ends of phrases? Did Korean subjects imagine a fermata at the end of each phrase? Or, were they hesitating and taking the time to decide whether to come in on a downbeat or an upbeat for the next phrase? What is even more interesting is that the number of beat(s) added at ends of phrases was different across individuals, resulting in different phrases that ended up reclaiming the missing anacrusis from phrase 1. This suggests that the metric perception of "Happy Birthday" is highly unstable among Koreans, and

it may be helpful to try to determine whether the act of clapping is confounding with the subjects' metric and rhythmic perception (as well as their production) of the song.

2.3.3 Power of Duple Preference to Manipulate Rhythm

Before moving on to discuss the results of native English speakers, it is worthwhile to make note of one Korean subject who managed to (quite ingeniously) transform the rhythm of the song in order to clap consistently in the 4/4 time signature with an anacrusis.



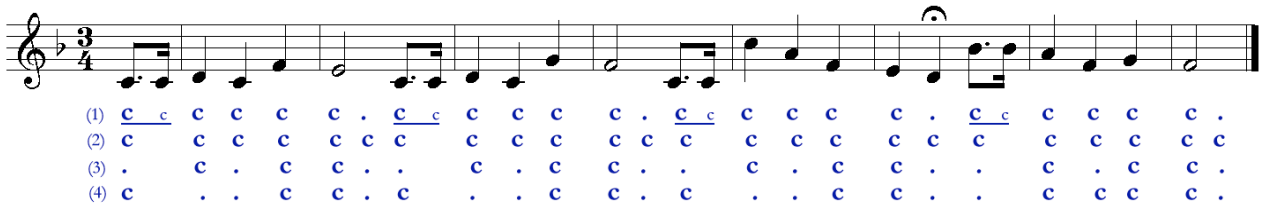
<FIGURE 12: Manipulated Rhythm>

The subject struggled at first to fit her clapping to the song, and ended up changing the song's rhythm in order to clap consistently across all the phrases, each beginning with an anacrusis. Interestingly, the subject was not aware that she had manipulated the rhythm of the original song. This phenomenon again suggests that the act of clapping has a potential to significantly bias the way singers perceive the song's metric structure (and therefore may be a suboptimal experimental method for measuring perception of meter).

2.3.4 Observation of Native English Speakers (with No Knowledge of Korean)[10]

Clapping patterns for native speakers of English were much more varied. **FIGURE 13** illustrates the performances of four native English speakers:

[10] Excluding one subject who clapped "correctly" according to the notated score.



<FIGURE 13: Clapping Results for Native English Subjects>

Except for one subject who clapped on every beat (#2), native speakers of English clapped based on syllabic emphasis, which naturally reinforced the melodic emphasis.[11] Because the claps were placed where there are syllabic emphases of the text, and because the lyrics for the four phrases are identical with a slight exception of phrase 3, the clapping patterns remained consistent *across* phrases. This is in direct contrast to performances by the Korean subjects, whose clap patterns were highly inconsistent across phrases as a result of extraneous beats added at ends of phrases.

Conversely, unlike the data from the Korean subjects that are characterized by a strong duple pattern, none of the clap patterns performed by the native English speakers display recurring 2- or 3-beat patterns. Consequently, we could not infer from the clap patterns the native English speakers' perceived metrical structure of the song.

2.3.5 A Note on Participants who Clapped Correctly

Two Korean and one native English speakers who clapped correctly (on beats that correspond to the downbeat in the notated score) commented that they felt sure that the way they clapped was "correct," although none of them claimed to have explicitly seen the score to "Happy Birthday" before. All three of the subjects have had extensive training in music.

[11] As the lyric for "Happy Birthday" was originally written in English, it makes sense that the linguistic accentuations and musical accentuations are aligned in a consistent manner.

2.4 DISCUSSION

2.4.1 Conclusion

As expected, a majority of Korean subjects clapped on the very first beat of "Happy Birthday," suggesting that they failed to perceive the anacrusis, possibly as a result of differences in syllabic emphases derived from the Korean lyrics. Similarly, native speakers of English tended to rely on the syllabic stress of English lyrics to determine their clapping patterns—which looked vastly different from that of Korean subjects.

This preliminary study unfortunately encountered various problems, and called for a follow-up study with a revised experimental method. The following section discusses these shortcomings and suggests a revised design.

2.4.2 Problems Encountered

Some of the major problems encountered in this experiment lie in the choice of subjects and the method used to measure subjects' metric perception of "Happy Birthday". The following describes a revised choice of subjects and equipment for a follow-up study:

<u>Subjects</u>

Rather than comparing Korean-English bilinguals with native speakers of English who cannot speak Korean, the revised experiment should make finer comparisons within Korean-English bilinguals such that the level of familiarity of the Korean language would serve as an independent variable. This will allow us to pick out the effects that native language and the level of fluency in the language have on metric perception.

Procedure & Apparatus

Instead of clapping on perceived strong beats, subjects should tap their finger along to the rhythm of "Happy Birthday" to an apparatus that can detect the intensity and duration of the perceived beat. A MIDI keyboard or a contact microphone, both of which can track and record tap timing and intensity, would be ideal for this purpose.

Clapping was problematic because (1) it made the subjects become conscious of how they were perceiving the beats, (2) it did not allow for detecting subtle differences the level of emphasis that the subjects perceived for each beat, and (3) it was a discrete (all-or-nothing) action that required too great of a motor movement from the performer with a blatant feedback of a loud sound. Using an apparatus such as a contact microphone would minimize these problems encountered with the method of clapping.

2.4.3 Separating out Two Factors: Meter & Anacrusis

Two inter-related but separable phenomenon have been observed among the native Korean speakers during this preliminary study. First is their duple clapping pattern to the triple-metered song. Second is their tendency to clap on the very first beat of the song, suggesting their failure to perceive the anacrusis.

Because metric perception is affected by so many factors—including cues in the musical line, linguistic accentuations, and possibly cultural upbringing—it is difficult to calculate the precise meter a singer perceives via indirect behavioral measurements. In fact, the inferred meter is likely to change throughout the song based on the interaction of various metric cues. Thus, for the follow-up experiment, we do not try to calculate the perceived meter, but rather, attempt to solve a smaller chunk of this problem by determining the contour of perceived beat-strengths.

Specifically, the revised experimental design examines the subjects' beat-strength perception at the beginning of each measure (according to the original notated score in 3/4) in order to determine the level of anacrusis perception. If the extent to which the anacrusis is perceived changes for different parts along the song, we then have a greater reason to believe that our subjective sense (or our brain's calculation) of the musical meter is a fluid process that depends on the interaction of various cues between the musical line and the linguistic text. And in the end, the data we obtain on anacrusis perception will further enlighten us on the bigger issue of metric perception.

CHAPTER III

A Follow-Up Experiment (2007-2008) :

Syllabic Intensity Tracking through Finger Tapping

3.1 REVISED QUESTION & HYPOTHESIS

This follow-up experiment aims at determining the extent to which subjects perceive the anacrusis in the original English version and in the Korean-translated version of "Happy Birthday." Again, the motivation behind this question lies in understanding how conflicting inferential metric cues in the text and the tune of a song get resolved. This revised question is also directly relevant to understanding the more complicated problem of how the contour of perceived beat strengths contribute to the overall metric perception of the song.

We hypothesize that the anacrusis will be more pronounced (1) when a native English speaker sings the song in English as opposed to when a native Korean speaker sings the song in English [across-group], and (2) when the song is sang in English as opposed to in Korean by a Korean-English bilingual [within-subject].

Clearly, this hypothesis has been formed based on the assumption that the linguistic accentuations from the lyrics interact with the musical accentuations derived from the tune; for otherwise, the language used to sing the song should not have an impact on the beat strength perception. Moreover, if it turns out that anacrusis is perceived when English lyrics are used but

not when Korean lyrics are used, we may conclude that prosodic accentuation can in fact dominate over musical accentuation.

3.2 METHOD: FINGER TAPPING

3.2.1 Subjects

Thirty students were recruited through an email advertisement to participate in a 30-minute study. Twelve native-English speakers who do not speak Korean (hereafter "Group I") and eighteen Korean-English bilinguals (hereafter "Group II") participated in the study. Group II was further categorized into the following four subgroups:

II_1: Korean-Americans whose native language is English and who are not very familiar with Korean (n=5)

II_2: Korean-English bilinguals who feel about equally competent in English and Korean. They were all students born in Korea who moved to the United States before high school. (n=5)

II_3: Korean-English bilinguals who feel more comfortable with Korean. They were all Korean students who came to the Untied States for college. (n=6)

II_4: Korean graduate students who did not come to the Untied States until after college, having spent approximately 25 or more years living in Korea. (n=2)

3.2.2 Equipment: Hardware & Software

(i) A USB condenser microphone was used in conjunction with computer1 to collect singing and speech samples. This data was recorded using *Audacity*.

(ii) An acoustic drum trigger, attached on the bottom surface of a table, was used to record finger tapping. A small blue sticker on the table surface directly above the contact microphone guided the subjects' placement of taps. The drum trigger was connected to a quarter-inch instrument cable, which served as an input to a MOTU UltraLite FireWire Audio Interface that was connected to computer2. Again, the tap data was recorded using *Audacity*. In addition, *Praat* was used to extract from the recordings quantitative values for tap intensity, which was then organized and visualized using *Microsoft Excel*.

Because the drum trigger was attached to the bottom of the desk surface and connected to the computer in an inconspicuous way, subjects in general did not realize that their taps were being recorded. (That is, during the tap trials involving simultaneous finger-tapping and singing, most subjects thought that it was their *singing* that was being recorded.) The lack of awareness of the contact microphone was beneficial because this allowed the subjects to tap without being conscious of the fact that their tapping was actually the measurement of interest for the study.

3.2.3 Experiment Session Procedure

Part 1: Short Questionnaire

A short questionnaire was used to associate the subject with a code number, and to determine the subject's native language, as well as any other language(s) in which the subject is able to sing the song "Happy Birthday".

Part 2: Perform Tasks in the Subject's Primary Language (hereafter "L1")[12]

[1] Sing the song "Happy Birthday" using the syllable /la/[13].

[12] Note: L1 is always English for Group I. L1 is either English or Korean for Group II

This step helped the subjects to naturally figure out the syllabic rhythm of the song.

[2] Tap index finger along on the syllables while singing the song using the syllable /la/.

This step allowed subjects to practice tapping on syllables

[3] Recite in normal speech the lyrics of the song.

This step was used to ensure the subjects' correct knowledge of lyrics, and to help them become aware of the linguistic features of the text without the music. The name "Sarah" was used in the English lyrics, and "Ji Eun" for the Korean lyrics.

[4] Sing the song "Happy Birthday" using the actual lyrics.

This step was in preparation for the next (ultimate) step, to allow subjects to feel comfortable singing the song.

[5] Tap index finger along on the syllables while singing the song using the actual lyrics.

This was the data of interest in our study. Five samples were collected.

Part 3[14]: Perform Tasks in the Subject's Secondary Language (hereafter "L2")

This part is identical to Part 2, except for the change in language.

Part 4: Survey

A final survey was used to collect basic information on the subjects' musical background, language background, and miscellaneous feedback.

---

[13] Note: the syllable /la/ , though somewhat different in quality, is available to both Korean and English phonetics.

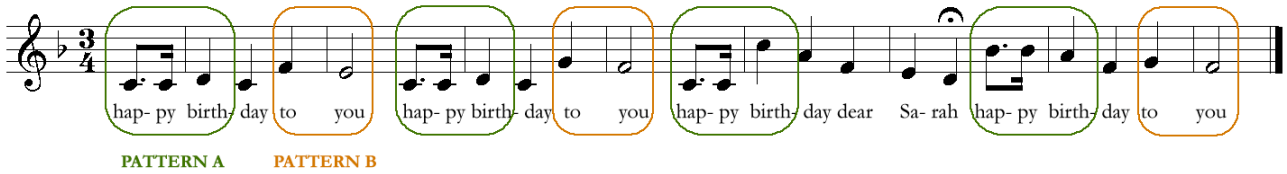
[14] *Part 3* was conducted for Group II only; L2 is either Korean or English, whichever language not used during *Part 2*. L2 is the subject's less-familiar language.

3.2.4 Analysis Procedure

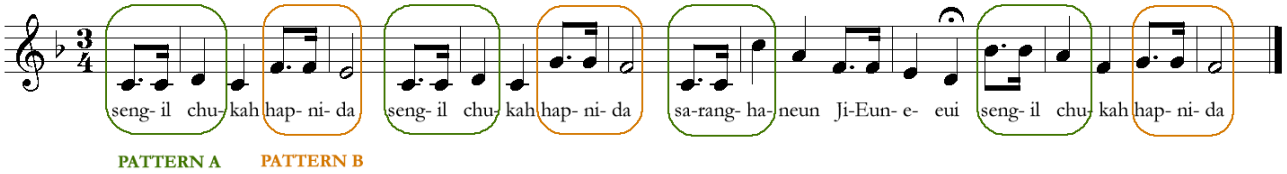
(1) The recordings of 5 trials of finger tapping (per subject-language) were made in *Audacity*, and saved as 5 individual .wav files.

(2) Using *Praat* tap-intensity measurements from each of the 5 subject-language .wav files were extracted and matched to the syllables in the text (equivalently, notes in the music). This data was then written to a spreadsheet file.

(3) Using *Excel*, we calculated the intensity ratio between two consecutive syllables. Data for *ha→ppy* followed by *ppy→ birth* (hereafter "Pattern A"), as well as data for *to→you* (hereafter "Pattern B") from each subject-language were extracted and compiled for further analysis.



<FIGURE 14: *Pattern A* and *Pattern B* for English version>



<FIGURE 15: *Pattern A* and *Pattern B* for Korean version >

(4) The change in tap intensity between the two beats that make up Pattern A and Pattern B were calculated (hereafter "RatioA" and RatioB", respectively) to infer the extent to which subjects perceived the first beat of the patterns as anacrusis.

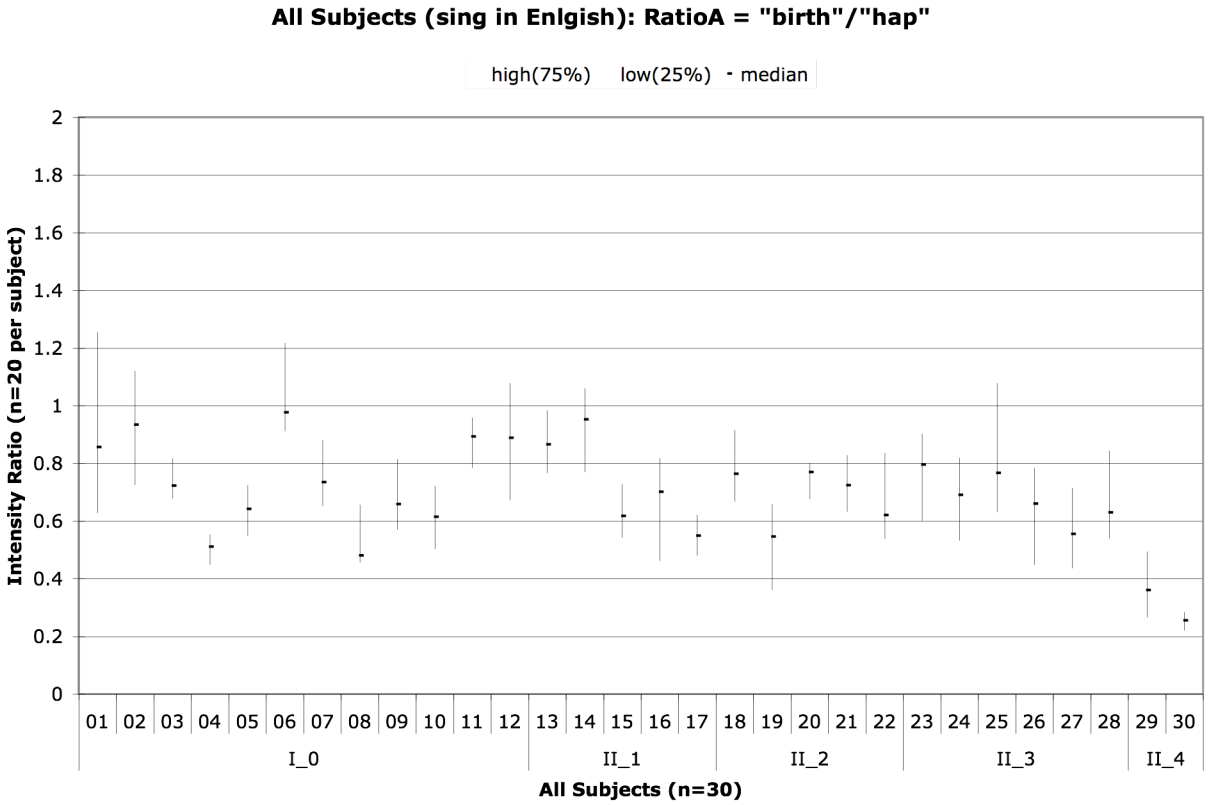
3.3 RESULTS

3.3.1 Across-subgroup Comparisons

<u>RatioA: singing in English</u> (RatioA$_E$)

GRAPH 16 illustrates the intensity ratio for *Pattern A* sung in English for all 30 subjects. RatioA$_E$ is calculated by the following equation: $RatioA_E = \left(\frac{"py"}{"hap"}\right)\left(\frac{"birth"}{"py"}\right) = \frac{"birth"}{"hap"}$, where "[*syllable*]" stands for the intensity with which the finger was tapped on the syllable.

For each of the 30 subjects, 20 RatioA$_E$'s were calculated: each singing trial consisted of 4 instances of *Pattern A* (see **FIGURE 14**), and we repeated the trial 5 times. We draw a line from the 25th to the 75th percentile of the 20 RatioA$_E$'s, with a tick mark at the median.



<GRAPH 16: RatioA$_E$ >

We interpret the value for $RatioA_E$ as follows:

- $RatioA_E$ =1: The syllable "hap" and "birth" were tapped at same intensity.
- $RatioA_E$ <1: The intensity decreased from "hap" to "birth." We infer that the first beat (corresponding to "happy") was perceived as the downbeat.
- $RatioA_E$ >1: The intensity increased from "hap" to "birth." We infer that the first beat (corresponding to "happy") was perceived as the anacrusis.

The across-subject intensity ratios for the 12 subjects in Group I are quite spread out, almost as much as the data for the 18 subjects in Group II.[15] This large spread within Group I may in part be caused by that, though all Group I subjects were native English speakers, many of them spoke other languages, including Mandarin, Swahili, Hebrew, and Spanish.[16] Also, because many factors beyond native language affect beat-strength perception, we must be careful in generalizing the subjects within the same group (or subgroup) as a single unit.

Nonetheless, the data still allows us to make across-group comparison for the 4 subgroups in Group II. We find a negative correlation between Group II's subgroup number and $RatioA_E$: as we go from Korean-Americans fluent in English (subgroup II_1) to foreign Korean students fluent in Korean (subgroup II_4), we find that $RatioA_E$ roughly decreases[17].

---

[15] The standard deviation of the median values of $RatioA_E$ for Group I (n=12) is 0.1664, and the standard deviation of the median values of $RatioA_E$ for Group II (n=18) is 0.1680.

[16] To try to minimize this effect in future studies, Group I should be restricted to native English speakers who have little knowledge of other languages.

[17] The regression line through four points, [(1, av(median $RatioA_E$ for subgroup II_1)), (2, av(median $RatioA_E$ for subgroup II_2)), (3, av(median $RatioA_E$ for subgroup II_3)), (4, av(median $RatioA_E$ for subgroup II_4))] is

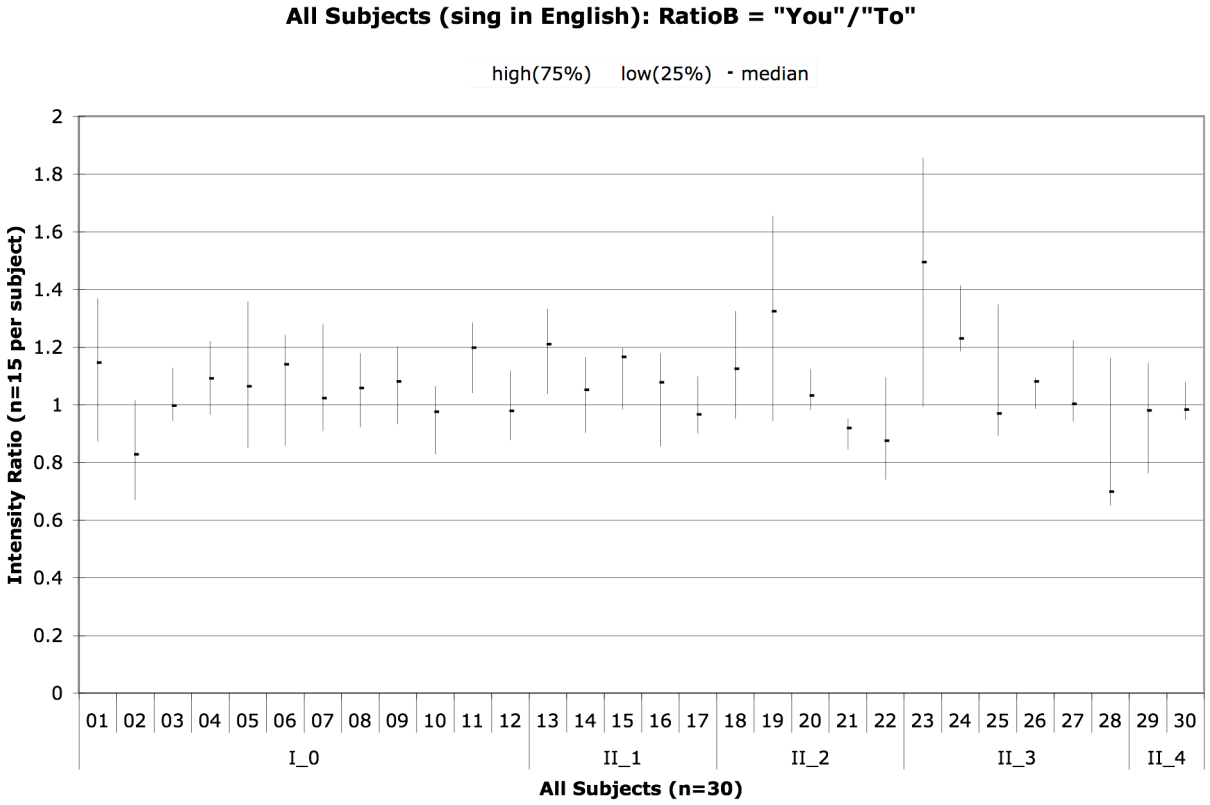
y=-0.1288x + 0.9258, with **r=-0.8385**.

Note the negative slope, but also realize that the magnitude of the slope is meaningless by itself because the x values were set arbitrarily as 1, 2, 3, and 4 according to the subgroup numbers; the meaning of the slope should be considered only insofar as making comparisons with the slope values for $RatioB_E$, $RatioA_K$, and $RatioB_K$.

This observation suggests that, as hypothesized, $RatioA_E$ decreases as the familiarity of the Korean language increases. In other words, the first beat of Pattern A is perceived stronger relative to the second beat for subjects who are more familiar with Korean.[18] This in turn suggests that greater familiarity of Korean is correlated with reduced perception of the anacrusis.

<u>RatioB: English</u> ($RatioB_E$)

GRAPH 17 illustrates the intensity ratio for *Pattern B* sung in English for all 30 subjects. $RatioB_E$ is calculated by the following equation: $RatioB_E = \frac{"you"}{"to"}$.



<GRAPH 17: $RatioB_E$ >

[18] Recruiting more participants in group II_4 would have allowed us to test the robustness of this subgroup's low $RatioA_E$ values.

For each of the 30 subjects, 15 $RatioB_E$'s were calculated: each English singing trial consisted of 3 instances of *Pattern B* (see **FIGURE 14**; the third phrase is excluded because the lyrics in it are different from the lyrics in phrases 1, 2, and 4. Furthermore a fermata at the end of phrase 3 alters its tempo and rhythm), and we repeated the trial 5 times. Again, the graph shows a line drawn from the 25th to the 75th percentile, with a tick mark at the median.

The interpretation of $RatioB_E$ is similar to that of $RatioA_E$: $RatioB_E > 1$ implies perception of the anacrusis in the English singing of Pattern B, while $RatioB_E < 1$ implies a lack of anacrusis perception in the English singing of Pattern B.

Surprisingly, in contrast to $RatioA_E$, to we do not observe a negative correlation between the subgroup number and $RatioB_E$: The slope of the regression model is very close to zero[19], which can be visualized in **GRAPH 17**.

In comparing **GRAPH 16** and **GRAPH 17**, we note that in general $RatioA_E < 1$ while $RatioB_E > 1$. This suggests that Pattern A has a different beat-strength contour than Pattern B, even though the notated score suggests the first beat of both patterns should be an anacrusis. Though the observation that $RatioA_E < 1$ is quite puzzling (and possible explanations to this finding is offered in *section 3.4.2*), the differences observed between $RatioA_E$ and $RatioB_E$ is actually not all that surprising given that Pattern A and Pattern B have different text, as well as contrasting rhythm and melodic lines, associated with them. A take-away point for this observation is the need perform analysis on Pattern A and Pattern B separately.

---

[19] The regression line through four points, [(1, av(median $RatioB_E$ for subgroup II_1)), (2, av(median $RatioB_E$ for subgroup II_2)), (3, av(median $RatioB_E$ for subgroup II_3)), (4, av(median $RatioB_E$ for subgroup II_4))] is

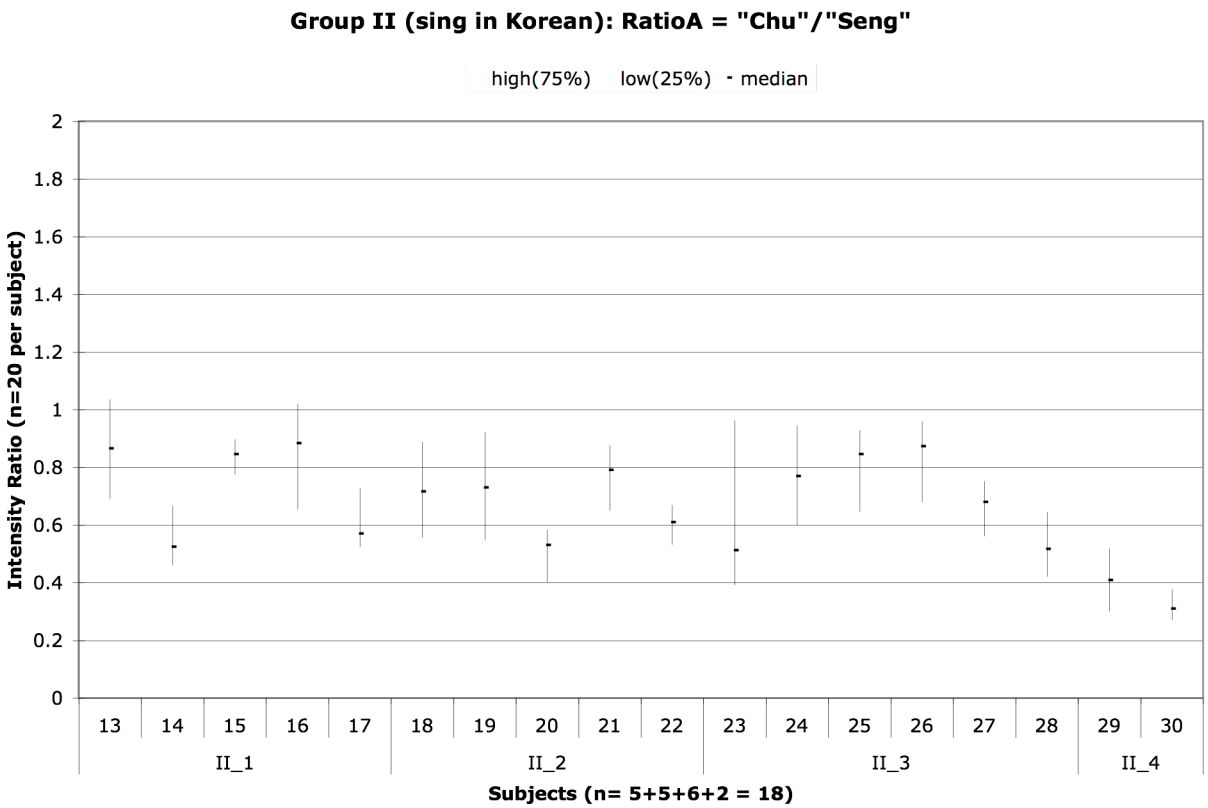
y=-0.0315x + 1.1316, with **r=-0.8115**.

Note the slope is close to 0, especially when compared with the slope of the regression calculated for $RatioA_E$.

<u>RatioA: Korean</u> ($RatioA_K$)

Similar to in the case for English, $RatioA_K$ was calculated for the 18 Group II subjects using the following formula: $RatioA_K = \left(\frac{"il"}{"seng"}\right)\left(\frac{"chu"}{"il"}\right) = \frac{"chu"}{"seng"}$.

The placements of syllables *seng*, *il*, and *chu* corresponds to the English syllables *hap*, *py*, and *birth*, respectively.



<GRAPH 18: $RatioA_K$ >

We confirm through this graph that $RatioA_K<1$, which means that the first beat of Pattern A was tapped harder than the second beat among Korean subjects. This is consistent with our hypothesis that the misalignment of the stress patterns in text and music for the Korean translated version would reduce the likelihood of perceiving the anacrusis notated in music.

The two subjects in the subgroup II_4 are characterized by the two lowest $RatioA_K$ values. We also find a slightly negative-sloped intensity-ratio vs. subgroup tendency[20]. Both of these observations are similar to that made for $RatioA_E$ previously.

RatioB: Korean ($RatioB_K$)

Finally, $RatioB_K$ was calculated for the 18 Group II subjects using the following formula:

$$RatioB_K = \left(\frac{"ni"}{"hap"}\right)\left(\frac{"da"}{"ni"}\right) = \frac{"da"}{"hap"}.$$

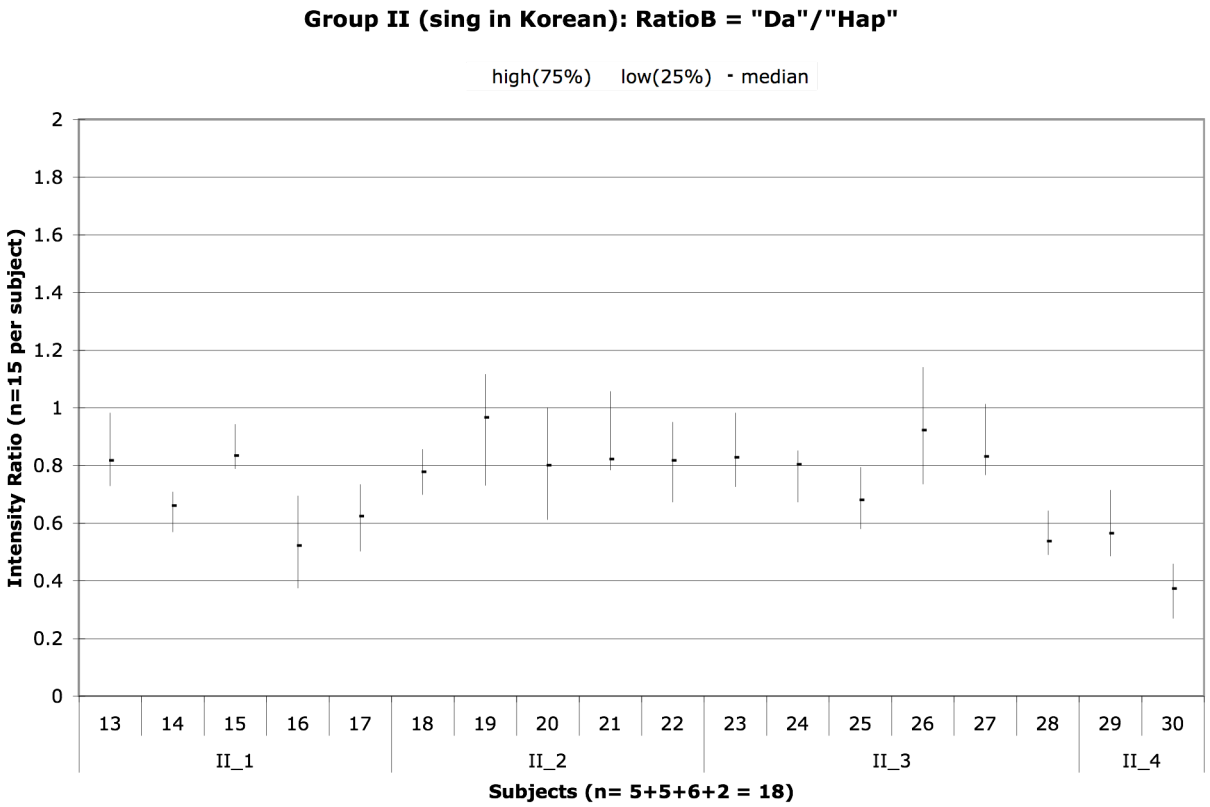
The placement of *hap-ni* and *da* corresponds to the English syllables *to* and *you* respectively. Here, note that $RatioB_K$ for Korean is calculated in a manner similar to $RatioA_E$ and $RatioA_K$—as a product of two ratios—in that there are three syllables that occur over the two-beat period. The rhythm in which these syllables occur in Pattern B is the same as the rhythm for Pattern A (refer to **FIGURE 14** and **FIGURE 15** for clarification).

We confirm in **GRAPH 19** that $RatioB_K<1$ as expected, indicating that Group II subjects tap the first beat of Pattern B stronger than the second beat. Again, the two subjects in the subgroup II_4 are characterized by having the two lowest $RatioB_K$ values, although we actually do not find a strong negative-sloped "intensity-ratio vs. subgroup" tendency in $RatioB_K$ as we did for $RatioA_E$ (**GRAPH 16**) and $RatioA_K$ (**GRAPH 18**) [21].

---

[20] The regression line through four points, [(1, av(median $RatioA_K$ for subgroup II_1)), (2, av(median $RatioA_K$ for subgroup II_2)), (3, av(median $RatioA_K$ for subgroup II_3)), (4, av(median $RatioA_K$ for subgroup II_4))] is
y=-0.1111x + 0.8968, with **r=-0.8232**.
Note the negative slope, but also realize that the magnitude of the slope is meaningless by itself because the x values were set arbitrarily as 1, 2, 3, and 4 according to the subgroup numbers; the meaning of the slope should be considered only insofar as making comparisons with the slope values for $RatioB_E$, $RatioB_{E,}$, and $RatioB_K$.

[21] The regression line through four points, [(1, av(median $RatioB_K$ for subgroup II_1)), (2, av(median $RatioB_K$ for subgroup II_2)), (3, av(median $RatioB_K$ for subgroup II_3)), (4, av(median $RatioB_K$ for subgroup II_4))] is
y=-0.0738x + 0.8756, with **r=-0.5972**.
Note that while the slope is negative, the magnitude is less than that of $RatioA_E$ and $RatioA_K$.



<GRAPH 19: RatioB$_K$ >

3.3.2 Within-Subject Comparisons (Group II only: English vs. Korean)

Rationale

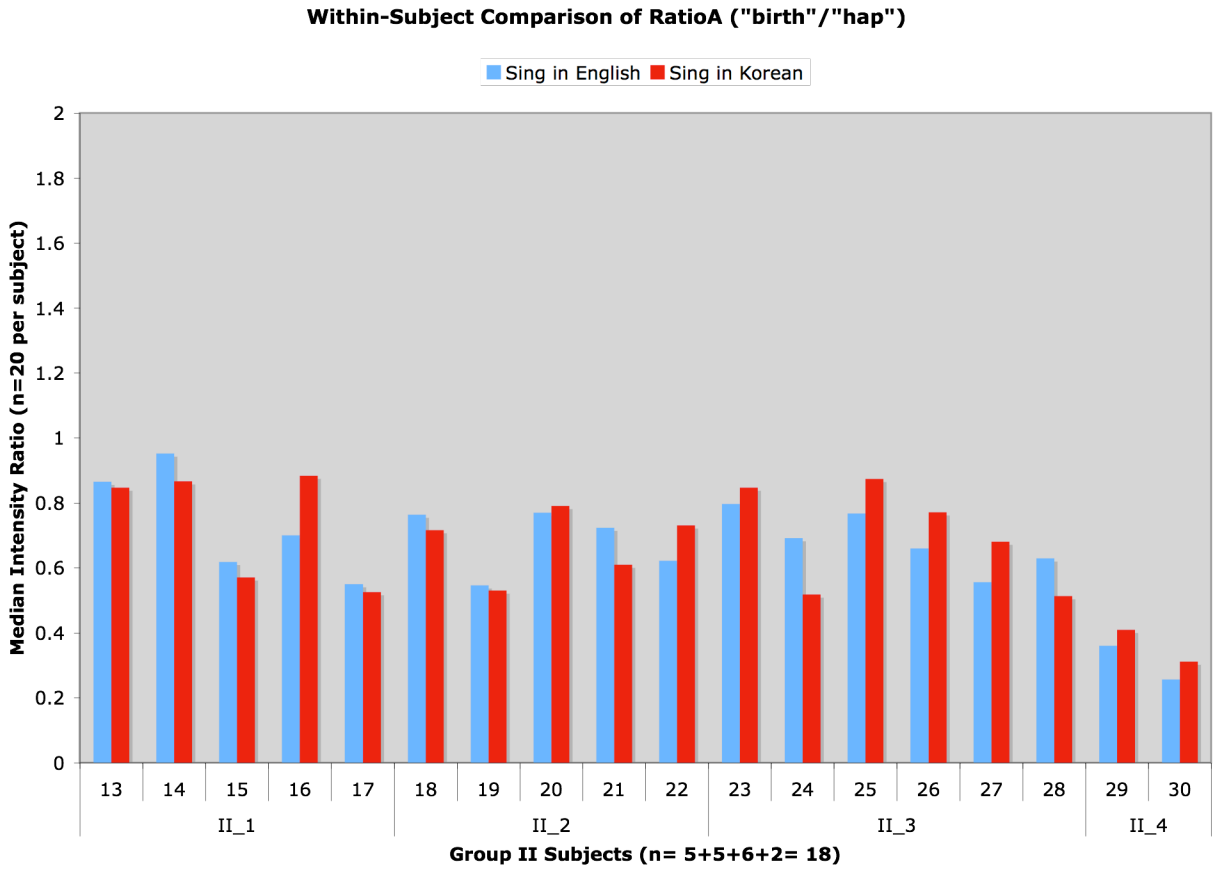
As pointed out previously, because our intensity-ratio data frequently showed a large variability within a subgroup, we had to be cautious in generalizing a subgroup's tendency by coming up with some kind of a characteristic accentuation-ratio number for each of the subgroups. Consequently, without a clean way of seeing a subgroup as a single unit, our across-subgroup analysis in *section 3.3.1* was not as effective as we had hoped. Thus, we now try a comparison on how the language affects beat-strength perception *within* an individual. If we consistently find a significant difference in Pattern A (between RatioA$_E$ and RatioA$_K$), or in

Pattern B (between $RatioB_E$ and $RatioB_K$) within an individual, we may then be able to conclude on the online effects of language on beat-strength perception of songs.

Pattern A: $RatioA_E$ vs. $RatioA_K$

For each individuals within Group II, we compare side-by-side the median $RatioA_E$ (blue) with the median $RatioA_K$ (red).

We observe in **GRAPH 20** that Pattern A does not show any online effects of language: For the 18 subjects in Group II, the difference between a given subject's $RatioA_E$ and $RatioA_K$ is insignificant[22]. Refer to *Appendix 5.6* for the numerical data and calculations.



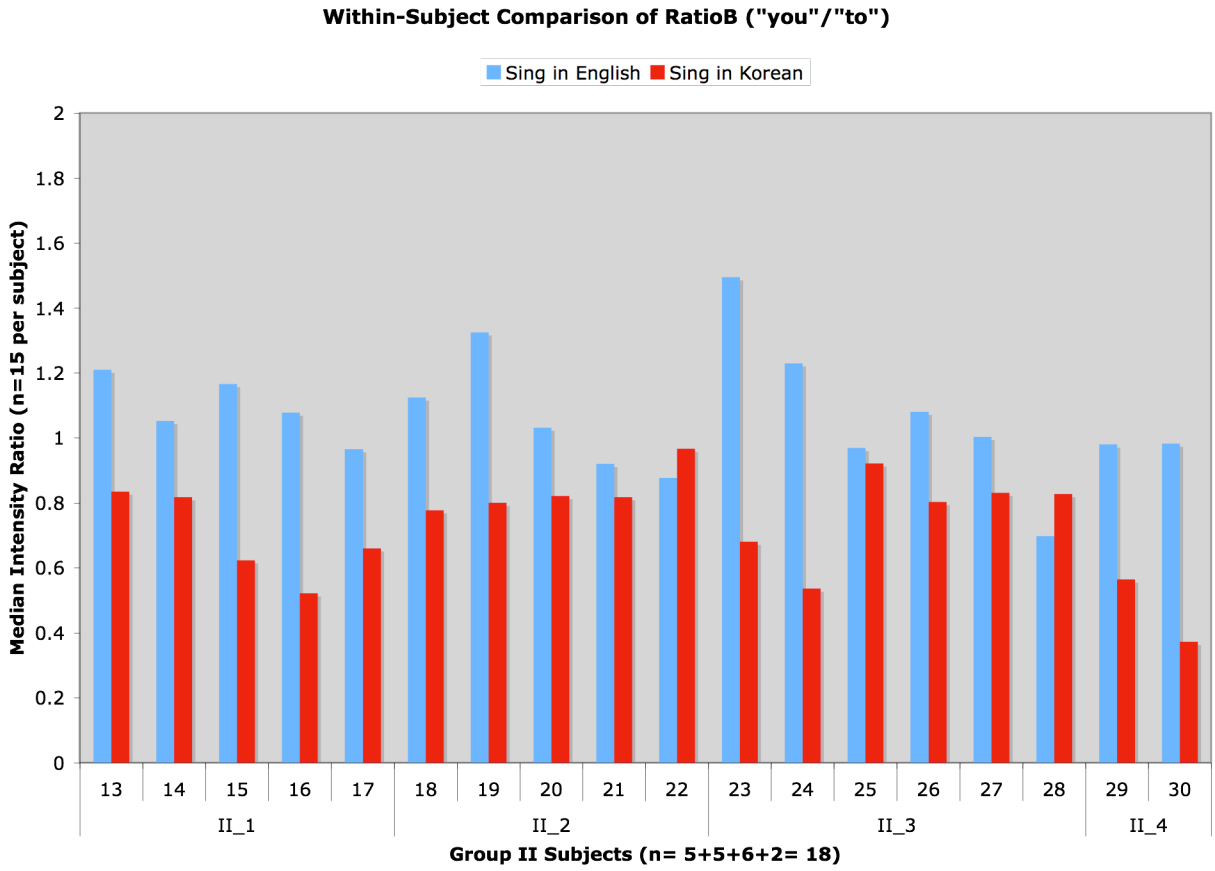
<GRAPH 20: RatioA, English vs. Korean>

[22] For the 18 $RatioA_E - RatioA_K$ values calculated for each subject in Group II, we find that the difference ranges from **-0.1834 to 0.1733**, with an **average of -0.0092** and **standard deviation of 0.0973**. (See Appendix 5.6)

Pattern B: $RatioB_E$ vs. $RatioB_K$

For each individuals within Group II, we now compare side-by-side the median $RatioB_E$ (blue) with the median $RatioB_K$ (red).

Even though Pattern A did not show online effects of language, we were surprised to see a dramatic change in tapping for Pattern B based on the language used to sing. As conveyed by **GRAPH 21**, the general tendency is that $RatioB_E > 1$, suggesting a definite perception of the anacrusis when the song is sung in English. In contrast, $RatioB_K < 1$, suggesting that the anacrusis was not perceived when sung in Korean.



<GRAPH 21: RatioB, English vs. Korean>

Furthermore, 16 out of 18 subjects in Group II are characterized by having a larger $RatioB_E$ than $RatioB_K$, suggesting that singing in English consistently heightened the subject's perception of the anacrusis in Pattern B. In fact, we calculate the difference between $RatioB_E$ and $RatioB_K$ for the 18 subjects, and find that the difference is quite significant[23]. Refer to *Appendix 5.6* for the numerical data and calculations.

Possible explanations as to why Pattern B displays such strong online effects of language while Pattern A does not are offered in *section 3.4.1* and *section 3.4.2*.

3.4 DISCUSSION

According to our across-subgroup analysis, we observed that the familiarity of Korean and the perception of anacrusis tended to be roughly negatively correlated. A strong factor in motivating this result was the data from subgroup II_4, which almost always had the lowest intensity-ratio values. Even though we only had two subjects in the subgroup II_4 and larger sample size for this subgroup is needed to verify this tendency, this observation is consistent with the original hypothesis that native Korean speakers fluent in Korean are less likely to perceive the anacrusis in "Happy Birthday" than those who are more familiar with English.

Furthermore, within-subject comparisons allowed us to find a significant change in the perceived beat-strength contour for Pattern B (but not for Pattern A) with the switch in the language used to sing "Happy Birthday". In the following subsections, we discuss possible reasons for why the linguistic effects were so prominent in Pattern B but not in Pattern A.

---

[23] For the 18 $RatioB_E - RatioB_K$ values calculated for each subject in Group II, we find that the difference ranges from **-0.1301 to 0.8134**, with an **average of 0.3339** and **standard deviation of 0.2612**. (See Appendix 5.6)

3.4.1 Rationale for online-effects of language for Pattern B

The most likely explanation for why Pattern B was characterized by such a significant difference between $RatioB_E$ and $RatioB_K$ has to do with the large disparity in the prosodic contour between the English text and the Korean text. The English contour articulate "to **YOU**", in comparison to the Korean contour of "**HAP-**ni da." Thus, we see that the weakest syllable in the English text occurs on "to", which corresponds to the strongest syllable in the Korean text, "hap". Conversely, the weakest syllable in the Korean text occurs on "da," which corresponds to the strongest syllable in the English text, "you". In this manner, the accentuation contour in the linguistic setting is completely the opposite for English and Korean in Pattern B.

Given this contrasting prosodic contour, we note that the English contour is consistent with the accentuations implied by the melody because "to" corresponds to the anacrusis that leads up to the strong beat on "you". Thus, the linguistic accentuation cues serves to reinforce this accentuation pattern. On the other hand, the Korean contour—which was shown to be the opposite of the English—conflicts with the musical accentuation, and the result is that, interestingly enough, the accentuation patterns in the Korean text override that of the music, generating a sense of a downbeat on what should have otherwise been an anacrusis according to the musical features and the notated score.

3.4.2 Unexpected lack of anacrusis perception for Pattern A

Now we offer several possible explanations as to why Pattern A, unlike Pattern B, failed to show a significant difference between $RatioA_E$, and $RatioA_K$. If we compare the linguistic accentuation contours between the English text (HAPpy-**BIRTH**) and the Korean text (**SENG**il-

CHU) occurring over Pattern A, we find that the contours are quite similar. The weakest syllable in English ("py") actually matches with the weakest syllable in Korean ("il"), and it is only the position of the strongest and the medium syllables ("birth") and ("hap") respectively, in English, and ("seng") and ("chu") respectively, in Korean, that are switched. Thus, the prosodic contour between English and Korean do not show as great of a contrast in Pattern A.

Moreover, the relative syllabic-emphasis difference between "hap" and "birth" in English is not great enough to fully reinforce the upbeat pattern found in the notated score. Similarly, the syllabic-emphasis difference between "seng" and "chu" in Korean is too subtle to fully overcome the existing feeling of anacrusis that comes with the notated score.

These factors explain why there was not a significant within-subject language differences in tapping Pattern A. But we still need an explanation as to why $RatioA_E$ was shown to be less than 1, despite the accentual implication of both. A possible explanation for this is the tendency for subjects, in performing the finger taps, to place a greater attack—a kind of an accent—on the very first beat of each phrases.

If this indeed turns out to be the case, the first beat of $RatioA_E$ (corresponding to the syllable "hap") would be tapped with greater intensity because it corresponds to the very first note of the phrases, while the first beat of $RatioB_E$ (corresponding to the syllable "to") is not applicable to this situation because Pattern B occurs in the middle of a phrases. Consequently, we are left with a smaller intensity-ratio data for a two-beat anacrusis patterns that occur in the beginning of phrases (such as Pattern A) than for those that occur elsewhere. Of course, this conjecture should be tested out in a future study.

3.4.3 Conclusion

This experiment, particularly the within-subject data over Pattern B, presents empirical evidence supporting the hypothesis that the Korean translated version of “Happy Birthday” is characterized by a different perceptual placement of strong beats than the original English version. Interestingly, the conflicting misalignment of accentual patterns between the text and the music is resolved by a change in the song’s metric structure (by shifting of a beat to eliminate the anacrusis), a rare phenomenon in which the metric structure of music becomes distorted in order to preserve prosodic accentuation.

Acknowledgments

I would like to express sincere thanks to the following people who made this project possible:

My thesis advisor, Professor Jonathan Berger, for introducing me to the field of music cognition and perception. His presentation of the topic on metric ambiguity, in particular, gave rise to my project. I am extremely grateful for his guidance and support over the past three years.

Professor Lera Boroditsky for her fascinating lectures and research on “language and thought”, which provided me with excellent experimental paradigms to work with. I am thankful for her insightful advice as my Second Reader.

Professor Chris Chafe and students in Music 220c, Spring Quarter 2006-2007, for helping me determine the optimal software and hardware setup for designing my follow-up study.

Todd Davies for his thoughtful advice during the Symbolic Systems Junior Honors Seminar and the Senior Honors Seminar.

And finally, the participants of my experiments for their time and cooperation.

CHAPTER IV

Bibliography

Addo, A.O. (1996). A multimedia analysis of selected Ghanaian children's play songs. *Bulletin of the Council for Research in Music Education*, 129, 1-28.

Baryshev, V. V. (1989). Speech Melody Perception: some specific features. *Soviet Journal of Psychology*, 10(2), 85-96.

Berger, J. & Berger, T. (2004). Monophonic polymeter. *Proceedings of the International Conference on Interdisciplinary Musicology (CIM04)*, Graz, Austria, April 2004.

Berger, J. & Gang, D. (1998). A Computational Model of Meter Cognition During the Audition of Functional Tonal Music: Modeling A-Priori Bias in Meter Cognition. *Proceedings of the International Computer Music Association,* Michigan.

Bergeson, T. R. & Trehub, S. E. (2005). Infants' Perception of Rhythmic Patterns. *Music Perception, 23*, 345-360.

Blacking, J. (1967). *Venda children's songs*. Johanessburg: Witwatersrand University Press.

Bolton, T. L. (1894). Rhythm. *American Journal of Psychology, 6*, 145-238).

Chen-Haftech, L. (1996). Effects of the pitch relationship between text and melody in Cantonese songs on young children's singing. Unpublished doctoral dissertation, University of Reading, UK.

Chen-Haftech, L. (1997). Music and Language Development in Early Childhood: Integrating Past Research in the Two Domains. *Early Child Development and Care*, 130 (1), 85-97.

Clarke, E. F. (1999). Rhythm and timing in music. In *The Psychology of music*, ed. Diana Deutsch: 473-500. San Diego: Academic Press.

Dell, F. & Halle, J. (2005). Comparing musical textsetting in French and in English songs to appear in proceedings volume from the first international Paris conference on the typology of metrical forms.

Dowling, W.J. (1988). Tonal structure and children's early learning of music. In J. Sloboda (ed.). *Generative Processes in Music*, 113-128. Oxford: OUP.

Fassbender, C. (1996). Infants' auditory sensitivity towards acoustic parameters of speech and music. In I. Deliege and J. Sloboda (eds). *Musical beginnings: origins and development of musical competence*, 56-87. Oxford: OUP.

Hannon, E. E., Snyder, J. S., Eerola, T. & Krumhansl, C. L. (2004). The Role of Melodic and Temporal Cues in Perceiving Musical Meter. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 956-974.

Kalmar, M. & Balasko, G. (1987). 'Music mother tongue' and creativity in preschool children's melody improvisations. *Bulletin of the Council for Research in Music Education*, 91, 77-86.

Majid, A., Bowerman, M., Kita, S., Haun, D. B. M. & Levinson, S. C. L. (2004). Can language restructure cognition? The case of space. *Trends in Cognitive Science, 8*, 108-114.

Mang, E. (2007). Speech-song interface of Chinese speakers. *Music Education Research*, 9(1), 49-64.

McMullen, E., & Saffran, J. (2004). Music and Language: A Developmental Comparison. *Music Perception*, 21(3), 289-311.

Minami, Y. & Umezzawa, Y. (1990). The situation in which a child sings an original song. In J. Dobbs (ed). *Music education: facing the future*, 131-134. Christchurch, NZ: ISME.

Morgan, T. & Janda, R. (1989). Musically-conditioned stress shift in Spanish revisited: empirical verification and nonlinear analysis. In Carl Kirschner, and Janet Ann DeCesaris (eds.) *Studies in Romance Linguistics*, Selected Proceedings from the XVII Linguistic Symposium on Romance Languages. New Jersey: Rutgers University.

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

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, 6, 688-691.

Papousek, M. & Papousek, H. (1981). Musical elements in the infant's vocalization: their significance for communication, cognition, and creativity. In L.P. Lipsitt (ed.). *Advances in infancy research*, 163-224. Norwood, N.J.: Ablex Pub. Co.

Ross, J. (2003). Same words performed spoken and sung: An acoustic comparison. *Proceedings of the 5th Triennial ESCOM Conference*, Hanover University, Germany.

Trehub, S.E. & Trainor, L.J. (1993). Listening strategies and infancy: the roots of music and language development. In S. McAdams and E. Bigand (eds.). *Thinking in sound: the cognitive psychology of human audition*, 278-327. Oxford: OUP.

Wright, M. (2007). "Survey of Models of Musical Rhythm" (invited talk, abstract only, full paper in print) *2nd Symposium on Music, Rhythm and the Brain*, Stanford, CA: Center for Computer Research in Music and Acoustics (CCRMA).

CHAPTER V

Appendix: Numerical Data

Note for the electronic (pdf) version:

The Appendix Content is in a separate Excel file, JieunOh_Honors_Appendix.xls