Too Cute for Words: An Investigation of Prosody’s Role in the Construction of Aegyo as a Speech Style

Shelby Strong
Washington University in St Louis

WUURD, the Washington University Undergraduate Research Digest, is published by the Office of Undergraduate Research once a semester each academic year. Applications for submission and Statement of Editorial Policy may be found online.

Follow this and additional works at: http://openscholarship.wustl.edu/vol8_iss1

Recommended Citation
Strong, Shelby, "Too Cute for Words: An Investigation of Prosody's Role in the Construction of Aegyo as a Speech Style" (2012). Washington University Undergraduate Research Digest, Volume 8, Issue 1.
http://openscholarship.wustl.edu/vol8_iss1/184

This publication is brought to you for free and open access by the Office of Undergraduate Research through Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu
Too Cute For Words: An Investigation of Prosody’s Role in the Construction of Aegyo as a Speech Style

Author: Shelby Strong

Shelby is a senior majoring in linguistics and East Asian studies. In Amy Baily’s Freshman Writing I course, she researched representations of gender in the media for a paper on the anime heroine Sailor Moon. This experience fueled her passion for research topics such as equality, discrimination, power, gender, and sexuality. Since then, her classes have provided her with a wider range of tools and perspectives for the exploration of these interests, and have led her to conduct this study as part of her Senior Honors Thesis in linguistics.

KEY TERMS

• Aegyo
• Korean Tones and Break Indices (K-ToBI)

ABSTRACT

Research suggests that prosody can be used by speakers to convey a variety of information to listeners. The present study attempts to investigate how prosody might play a role in the construction of aegyo, a Korean term for a type of charming, cute behavior that may refer to both linguistic and non-linguistic behavior. The focus of this study was the speech style which uses linguistic features stereotyped as conveying cuteness or charm. To test the hypothesis that there is a significant prosodic difference between aegyo and non-aegyo, two experiments were conducted. In Experiment 1, subjects read aloud passages in Korean that were constructed to elicit contrastive speech performances. The subjects’ performance was expected to exhibit more or less aegyo depending on the context information embedded in the passages. While there was a significant difference between the mean pitch of utterances for two out of the three subjects, there was no significant difference between the pitch range and mean absolute slopes of contrastive utterances. Additionally, analyses revealed that seven of the nine widely recognized K-ToBI IP-final boundary tones occurred in both aegyo and non-aegyo utterances. In Experiment 2, two judges listened to and evaluated the clips of utterances derived from recordings made in Experiment 1. Despite some subjectivity in judging, there was a high level of concordance between the judges. The study concluded that there was some evidence, at least in the case of mean pitch, suggesting a prosodic difference between aegyo and non-aegyo speech. Due to the fact that very few subjects participated, it was concluded that the results could not be used to make inferences about other populations and that the study should be replicated in the future with a larger number of subjects.

FACULTY MENTOR: BRETT KESSLER, PH.D.
ASSOCIATE PROFESSOR OF PSYCHOLOGY AND PNP

Professor Kessler currently studies psycholinguistics of reading and spelling. He is particularly interested in computational and statistical approaches to language, especially in the fields of phonology, historical linguistics, and the lexicon. His research has also explored how to statistically test the historical connections between languages.

ACKNOWLEDGEMENTS

I thank the Office of Undergraduate Research for support of this project. I would like to express my deep gratitude to Dr. Brett Kessler for his invaluable advising and to Dr. Mitchell Sommers for allowing me to use a sound booth in his lab. I would also like to thank the professors who I have been fortunate to have as teachers in classes that have instilled in me the desire to pursue research and who have encouraged me to ask questions and to seek answers. In particular I would like to thank Amy Baily whose teaching freshman year and continued support these past three years have inspired and encouraged me to pursue my interest in researching aegyo.

Peer Editor: Sarah Cohen, a senior majoring in English Literature and Psychology
INTRODUCTION

Aegyo is a Korean term for a type of charming, cute behavior frequently observed in contemporary Korean media including dramas, variety shows, reality shows, and televised music programs that feature appearances by Korean celebrity pop singers and idol groups as guests. Aegyo appears to be a type of performance that typically consists of either physical behavior or linguistic behavior or a combination of both that is perceived as cute. I have observed that in performances on television, females appear to be requested to perform aegyo more frequently than men and requests for aegyo are made of younger celebrities more frequently than older ones.

Despite aegyo’s seemingly pervasive presence in contemporary Korean media and despite the recent increases in the body of literature pertaining to Korean studies, only a few scholarly articles are written in English and, more often than not, only briefly mention aegyo. Therefore, I look to other research in the field of sociolinguistics, in particular speech styles and code-switching, as starting points for understanding aegyo. My observation of changes in pitch lead me to question the role that prosody plays in aegyo. The results of this study will be informative to a variety of scholars and non-scholars and will help establish a working definition of aegyo based on empirical evidence. This definition will be based on some of the linguistic features that often co-occur with speech perceived to be spoken in an aegyo speech style. The study will not answer questions of what aegyo “does”, (e.g. What are the motivations of speakers who employ aegyo? What does the employment of aegyo help speakers to achieve in their interactions?), but rather attempts to describe what aegyo “is” in part.

In the Dictionary of the Standard Korean Language aegyo is defined as a “cute manner or attitude that is shown towards others”.1 This definition however does not provide an adequate description; not all cute behavior is aegyo. “Cute” and “aegyo” are neither synonymous nor do they hold equal weight in meaning and connotation depending upon the context.

Understanding and defining aegyo should not be marginalized. Inoue asserts that “Japanese women’s language” (onna kotoba or joseigo) is a “space of discourse—understood as a complex ensemble of practices, institutions, representations, and power—in which the Japanese woman is objectified, evaluated, studied, staged, and normalized through her imputed language use and is thus rendered a knowable and unified subject both to herself and to others”.2 Likewise, I argue that aegyo is a socially constructed category and a part of a larger social discourse in Korea with far-reaching and significant consequences throughout society. The ability to behave in a cute manner is not biologically determined, though the frequent association of aegyo with females may be an effort to propagate this falsehood. Aegyo is made visible and circulated whenever it is performed or a performance of it is referenced. Individuals’ knowledge of aegyo is subjective, something that is learned from their experiences. It should be understood that aegyo is not something objective that can be defined universally across contexts. However, various social factors, such as the reinforcement of certain types of behaviors with positive or negative reactions, allow a norm for what is considered stereotypical aegyo to be maintained.

The sociolinguist Coupland wrote that “conventionalized associations often link
social categories together, for example, masculinity and ethnicity. There are many ways of being and talking ‘like a man’ or ‘like a black person’, but in some socio-cultural contexts the formation ‘masculine’ or ‘manly’ can come to be defined in terms of blackness. I would assert that aegyo can be perceived as a speech style, stereotypical in nature, that speakers may switch to or from to make salient the fact that they are indexing different social groups during conversation and therefore communicating their stance on a variety of issues, including their identity. Through his research, Barrett demonstrated that African-American drag queens are able to challenge and resist mainstream ideologies concerning gender and sexuality by switching between speech styles. It may be the case that performances of aegyo and style-switching involving aegyo may reinforce or subvert gender and sexuality ideologies. One of the objectives of this study is to provide information on the linguistic features that might potentially mark the aegyo speech style, information that will be useful for studies of style-shifting involving aegyo.

To investigate whether prosody significantly varies between aegyo and non-aegyo speech, four different variables were examined; mean absolute pitch slope, pitch range, mean pitch and IP-final boundary tones. Erez Levon defines pitch as “a perceptual phenomenon, intrinsically related to, but distinct from the physical act of vocal fold vibration. When the vocal folds vibrate, they always do so at a regular periodic frequency (e.g., 100 times per second, annotated as 100 Hz)”.

In the present study data collected was also qualitatively annotated using the Korean Tones and Break Indices Conventions (K-ToBI) written by Sun-Ah Jun. Jun wrote “the intonational structure of the standard dialect (=Seoul) of Korean has two intonationally defined prosodic units: Intonation Phrase (IP) and Accentual Phrase (AP)…An IP is marked by a boundary tone (%) and final lengthening”. According to Jun, the boundary tone occurs in the IP-final syllable and conveys pragmatic meaning and information about the sentence type. I hypothesize that there may be a significant difference between the types of boundary tones that are realized in the IP-final syllables of aegyo utterances and non-aegyo utterances.
METHOD

Experiment 1

Participants
Three native Korean speakers, two females and one male, voluntarily participated in the experiment, and were current undergraduates or recent graduates of Washington University in St. Louis between the ages of 20 and 23. All participants had lived in the Republic of Korea for 12 or more years, attended school in Korea for seven or more years, and had studied abroad, including their time spent attending college in the United States. All had studied one or more foreign language, including English, for 10 or more years and all stated that they were relatively fluent in English. All of the subjects stated that they were fluent in the standard Seoul dialect. One subject stated that he also spoke Gyeongsang dialect (경상도 방언) and another subject stated that she also spoke Chungcheong dialect (충청도 방언).

Materials
The experiment was divided into seven different tasks and all segments took place in the same location; a quiet, private sound booth. Tasks 1 and 4 of the experiment consisted of the subjects reading a passage aloud while being recorded. In Tasks 2 and 5, subjects were asked to answer three questions regarding the passages they read in Tasks 1 and 4. The data collected from Tasks 3 and 6 ultimately was not used but consisted of asking the subjects to retell the passages from Tasks 1 and 4 while being recorded. In Task 7, the subjects were asked to complete a demographic questionnaire to gather information about their background. The tasks took place in sequential order, e.g. Task 1 before Task 2 before Task 3.

The instruments used to collect data in Task 1 and Task 4 consisted of two passages. The two passages contained two nearly identical paragraphs followed mostly by dialogue. Both passages contained target phrases that did not vary or varied very little in their orthographic presentation. The paragraphs in both passages provided background information about the characters that appeared in the passage and contextualized the setting. Each dialogue consisted of a phone call between a character and her boyfriend. The girlfriend in Task 1 is named Seong-hwe and the girlfriend in Task 4 is named Young-mi. In both passages Seong-hwe is repetitively described as someone who lacks aegyo and doesn’t display a cute side to others, without actually stating the word aegyo itself in the passage. In both passages, Young-mi is repetitively described as someone who loves cute things, behaves cutely to her boyfriend, and frequently acts in a cute manner. Through the insertion of descriptors throughout the passages, including the phone call dialogue, the contrast between the contexts and differences between each character’s stereotyped personality was reinforced. For example, in sentences that contained target utterances, Young-mi was described as speaking in a cute voice whereas Seong-hwe was not.

It was expected that utterances “said” by Seong-hwe would elicit less aegyo than
utterances “said” by Young-mi. See *Example 1* for samples of target sentences containing utterances forming contrastive pairs. In the passages, there were also utterances that were expected to have little or no contrast, either because the character voicing them did not change, or because they were identical sentences in both passages. From these sentences, five pairs were selected as control pairs.

The instruments used to collect data in Task 2 and Task 5 consisted of identical questionnaires that asked the participants to describe the personalities of Seong-hwe and Young-mi, and to summarize the plot of the passage. These two tasks were constructed to ascertain that the subjects understood the context information embedded in the passages. After the experiment, the responses were read and it was determined that all of the subjects had comprehended the passages. The instrument used to collect data in Task 7 was a questionnaire used to learn about participants’ demographic backgrounds.

**Example 1**

Examples of Different Paired Utterances Embedded in Task 1 and Task 4

*Excerpts from Task 1 and Task 4 Containing Identical Utterances Contextualized to Contrast*

<table>
<thead>
<tr>
<th>Task</th>
<th>Korean Orthography</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task 1</td>
<td>“오빠!” 성회는 말했다. “나도 아이스크림 먹고 싶어!”</td>
</tr>
<tr>
<td>Identical</td>
<td>Non-Aegyo Context</td>
</tr>
</tbody>
</table>

(English Translation: “Boyfriend!” Seong-hwe said in a child-like voice. “I want to eat ice cream too!”)

<table>
<thead>
<tr>
<th>Task 4</th>
<th>Korean Orthography</th>
</tr>
</thead>
<tbody>
<tr>
<td>“오빠!” 응미는 아이 같은 목소리로 말했다. “나도 아이스크림 먹고 싶어!”</td>
<td></td>
</tr>
<tr>
<td>Identical</td>
<td>Aegyo Context</td>
</tr>
</tbody>
</table>

(English Translation: “[Boyfriend!” Young-mi said in a child-like voice. “I want to eat ice cream too!”)

(Example continued on next page.)
Example 1 (continued)

Examples of Different Paired Utterances Embedded in Task 1 and Task 4

**Excerpts from Task 1 and Task 4 Containing Similar Utterances Contextualized to Contrast**

**Task 1**  
Korean Orthography

“혹시 저녁도 사주면 안돼?”  

<table>
<thead>
<tr>
<th>Identical</th>
<th>Non-Aegyo Context</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Similar</td>
<td></td>
</tr>
</tbody>
</table>

(English Translation: “Would it be possible for you to buy dinner as well?” Seong-hwe asked.)

**Task 4**  
Korean Orthography

“오빠 오빠 저녁도 사주면 안돼?” 음미는 귀엽게 물었다.

<table>
<thead>
<tr>
<th>Identical</th>
<th>Aegyo Context</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Similar</td>
<td></td>
</tr>
</tbody>
</table>

(English Translation: “Boyfriend boyfriend would it be possible for you to buy dinner as well?” Young-mi cutely asked.)

**Excerpts from Task 1 and Task 4 Containing Identical Utterances Not Contextualized to Contrast**

**Task 1/Task 4**  
성회와 음미는 중학교 부터 친구 중의 친구였다.

<table>
<thead>
<tr>
<th>Identical, Non-Aegyo Context</th>
</tr>
</thead>
</table>

(English Translation: Seong-hwe and Young-mi had been the best of friends since middle school.)

**Procedure**

Participants received oral instructs to read the passages in Tasks 1 and 4 aloud in an entertaining manner, as though they were reading to a friend or family member. They were given a printed copy of the passages and recorded individually in a soundproof sound booth. Subjects were individually recorded with a Shure PG27 multipurpose side-address condenser microphone connected via a M-Audio MobilePRE USB Bus-powered preamp and audio interface to a desktop computer. During the recording, the speech produced was monitored and saved in a program called Wavesurfer. The raw data was recorded as WAV format in mono 16 bits at a sampling rate of 4410 Hz and then later transferred to a flashdrive and computer for storage and analysis. Later, clips
were made from the recordings using Audacity, a free audio editor. The format of the raw data was preserved. After subjects completed reading a passage in Tasks 1 and 4, subjects were handed printed questionnaires in Tasks 2 and 5. Participants were instructed to write a short response to each question. In Task 7, participants were instructed to fill out the questionnaire and ask the researcher if they needed clarification regarding any questions.

**Experiment 2**

Experiment 2 consisted of two tasks. In Task 1, subjects listened to clips and made judgments about whether or not they perceived a clip to be aegyo. The clips used in this experiment were made from Experiment 1 recordings. Task 2 consisted of a demographic questionnaire. It was hypothesized that utterances “said” by Seong-hwe would be judged to be non-aegyo by participants and conversely, utterances “said” by Young-mi would be judged to be aegyo.

**Participants**

There were two subjects, one male native Korean speaker and one female non-native Korean speaker (myself). Neither of the subjects had participated in Experiment 1 and both were 21 years of age at the time of the study. The male subject spoke standard Seoul dialect, had attended school in South Korea for 12 years, and had been studying and living abroad in the United States for the past two years. I studied Korean as a foreign language at Washington University for three years, as well as through a study abroad experience in Seoul. It should be noted that the possibility for biases influencing my judgment was high, as I had already heard the clips before my participation in the experiment.

**Materials**

The instrument used to collect data in Task 1 consisted of a response sheet. For each of the clips, judges used the response sheet to rate whether they considered a clip to be aegyo. The options for responses were limited to three categories, “yes,” “no,” and “unsure.” “Yes” meant that subjects perceived a clip to be aegyo, “no” meant that subjects did not perceive a clip to be aegyo, “unsure” meant that subjects were undecided on whether or not a clip was aegyo. The instrument used to collect data in Task 2 was a demographic information questionnaire, the same one used for Task 7 in Experiment 1.

**Procedure**

The experiment took place in a soundproof booth. For Task 1, the stimuli were compiled into a computer playlist before the experiment. During the experiment, the stimuli were presented to the subjects with headphones. The judges were not trained in any way or given guidance on how to determine if a clip was aegyo. Subjects were instructed to listen to each clip and then circle the label on the response sheet that they thought best fit the clip. The subjects were informed that they could request to hear a clip again. In Task 2, subjects were instructed to fill a questionnaire out and to ask the researcher if they needed clarification regarding a question.
RESULTS & DISCUSSION

Experiment 1

The hypothesis that a significant difference between utterances produced in contrastive contexts, engineered to elicit less or more aegyo, as reflected in the mean absolute slope of the utterances was investigated. Data collected from Tasks 1 and 4 was analyzed using Praat, a speech analysis program, to find the mean absolute slope values for each utterance. Then the average of each subject’s mean absolute slope values in semitones (st) was calculated. The results did not match predictions. It was expected that Task 4 utterances would have a higher mean absolute slope on average than Task 1 utterances. However, with the exception of Subject 3, there was a decrease on average (see Figure 1). The significance of the differences between the average mean absolute slope of Task 1 and Task 4 clips was analyzed by paired t-tests conducted in R, a free software environment for statistical computing, which found that the p-values for all subjects were all above the cut-off p-value of 0.05. It was accepted that there was no significant difference between the average mean absolute slope in Task 1 and Task 4 for all subjects. It was concluded that the performances of the subjects did not vary significantly in terms of the average mean absolute slope of the utterances they produced even when they were voicing different characters.

![Average Mean Absolute Slope of Subjects By Task](image)

**Figure 1**

To test the hypothesis that there is a significant difference between aegyo and non-aegyo utterances as reflected in the pitch range of utterances, the pitch range in semitones of each clip produced by each subject was calculated using Praat. These ranges were than averaged by subject. The average pitch range of subject 1 in Task 1 was 13.08725 semitones (st) and 16.10125 st in Task 4. The average pitch range of Subject
2 was 20.76 st and 21.02025 st in Task 1 and Task 4, respectively. The average pitch range of Subject 3 was 13.112 st and 17.41 st in Task 1 and Task 4, respectively. Utterances from Task 4 that were constructed to elicit “more aegyo” had greater pitch ranges on average for all subjects than those utterances analyzed from Task 1 (see Figure 2). This result was expected. However, for all subjects, the significance of the result was not supported by a paired t-test. The p-value of Subject 2, 0.4636, was deemed too high and unacceptable. The p-values for Subject 1 and Subject 3, .08242 and 0.06463 respectively, were much closer to 0.05 but still not low enough to be accepted as strong support for the existence of a significant difference in the average pitch ranges of Task 1 and Task 4 utterances.

![Average Pitch Range of Subjects By Task](image)

**Figure 2**

To evaluate the hypothesis that there is a significant difference between aegyo and non-aegyo utterances reflected in the mean average pitch of utterances, the mean pitch above 100 hertz of each clip was determined using Praat. The average of all of the clips’ mean pitches in Task 1 and Task 4 for each subject was then calculated. The result was that there was a higher average mean pitch for all subjects in Task 4 than Task 1 (see Figure 3). Subject 1 averaged a mean pitch of 4.38 st in Task 1 and 6.21 st in Task 4. Subject 2 averaged a mean pitch of 18.41 st in Task 1 and 19.05 st in Task 4. Subject 3 averaged a mean pitch of 15.88 st in Task 1 and 16.84 st in Task 4. A paired t-test was conducted and it was found that the p-values for Subject 1 and Subject 3 were 0.003375 and 0.003482, respectively, and were beneath the cutoff p-value of .05. Therefore the t-test for Subject 1 and Subject 3 supported the original hypothesis that the average mean pitch of speech differs significantly depending on whether the context of the utterances read by subjects was constructed to elicit more or less aegyo in Task 4 and Task 1. The conduction of a paired t-test for subject 2 resulted in a p-value of 0.2877, unacceptable support for a statistical difference.
Figure 3

The fourth hypothesis was that there is a significant difference between less aegyo and more aegyo utterances that can be seen in the difference between qualitative measurements made using the Korean Tones and Break Indices (K-ToBI) transcription convention. To test this hypothesis, clips were first analyzed using Praat. For each utterance, the Intonation Phrase (IP)-final word was divided into its respective syllables. Then the pitch contour of the final syllable of each IP-final word, which is known as the IP-final boundary tone, was coded for using the definitions for IP-final boundary tones stated in the K-ToBI transcription convention manual written by Sun-ah Jun (see Figure 4 for an example of the Praat display view used during coding.)

Figure 4
An Example Screenshot of the Display in Praat Used During Labeling of Utterance Boundary Tones (Clip 5Y)
In total there were 102 boundary tones coded for using K-ToBI. If we exclude the “control” pairs AN, BO, CP, DQ, and ER from the count then there were 72 boundary tones total; for each task, the boundary tones of 36 utterances were examined. It was expected that the utterances derived from clips A, B, N, and O would be rated in Experiment 2 as “aegyo” for both Tasks. Hence the boundary tones from the utterances taken from Task 4 clips and the utterances from clips A, B, N, and O were categorized as Aegyo. There were a total of 48 non-aegyo boundary tones whose IP-final boundary tones were examined and labeled. Pair CP and ER were declaratives and DQ was an interrogative. As none of the clips in Pairs CP, DQ, and ER were expected to elicit aegyo, they are included in the category Non-Aegyo in Table 1 below. There were a total of 54 Non-Aegyo utterances coded for. All nine of the boundary tones that have been classified in previous research, L%, H%, LH%, HL%, LHL%, HHL%, HLH%, HLHL%, LHLHL%, were observed in the examined utterances.

<table>
<thead>
<tr>
<th>IP-Final Boundary Tone</th>
<th>Task 1</th>
<th>Task 4</th>
<th>Non-Aegyo (Clips from Pairs CP, DQ, ER and Task 1 Clips)</th>
<th>Aegyo (Clips from Pairs AN, BO and Task 4 Clips)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L%</td>
<td>9</td>
<td>7</td>
<td>21</td>
<td>8</td>
</tr>
<tr>
<td>H%</td>
<td>10</td>
<td>7</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>LH%</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>HL%</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>LHL%</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>HLH%</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>LHLH%</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>HLHL%</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>LHLHL%</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Total Number of Labels</td>
<td>36</td>
<td>36</td>
<td>54</td>
<td>48</td>
</tr>
</tbody>
</table>

While the most frequently occurring IP-final boundary tones for both Task 1 and Task 4 utterances was L% and H%, the number of occurrences of the different boundary tones for Task 4 utterances were more closely bunched together in terms of their dispersion (see Figure 5). Two tones, HLHL% and LHLH%, which did occur in utterances from Task 4 did not occur at all in the Task 1 utterances. Except for these two tones, there were no boundary tones that were exclusive to a task or category. However, while these two tones only appeared in utterances that were expected to elicit aegyo, neither of these two tones were the mode for Task 4 or the category known as Aegyo. That is, HLHL% and LHLH% boundary tones were observed with less frequency in comparison to other boundary tones. This implies that these two tones are not strong indicators for utterances spoken in an aegyo speech style. Overall the utterances categorized as Task 4 and Aegyo were richer in diversity of IP-final boundary tone types than the utterances categorized as Task 1 and Non-Aegyo.
The boundary tones of utterances expected to be perceived as aegyo and non-aegyo were very similar in the types of IP-final boundary tones they contained. This result may imply that the current K-ToBI system is insufficient for measuring the prosodic difference between aegyo and non-aegyo utterances or that the prosodic difference between aegyo and non-aegyo is not significant from the perspective of the K-ToBI transcription conventions used. For this study a clear pattern between aegyo and non-aegyo utterances could not be ascertained using the K-ToBI. Therefore, the hypothesis that the K-ToBI transcription convention could be used to demonstrate a prosodic difference between less and more aegyo utterances was not supported. Of course, there is also the chance that the analysis approach using the K-ToBI transcription system may have allowed for greater error than other data analysis approaches due to the fact that the K-ToBI convention relies on researchers to make qualitative judgments based on their impression of the visible pitch contour. If there had been more than one person coding the clips, there may have been disagreement or alternate observations leading to different results. With that in mind, the fact that only one person participated in the coding of the utterances and that the number of utterances analyzed was fairly small and drawn from only three different speakers should be taken into consideration.

The question arises then: what does the data annotated using the K-ToBI demonstrate? While it is not reasonable to generalize the norm for larger populations of native Korean speakers from this experiment alone, it is amenable to discuss the possibilities that the results of the K-ToBI transcription invite future studies to explore and question. For one, is it possible that there are or would ever be IP-final boundary tones exclusive to one speech style, as exhibited in the data of this study? The idea that there is a boundary tone that is exclusive to aegyo seems unlikely. For example, a linguistic feature that marks a speech style may be used less frequently in favor of another feature that also marks the same speech style. Therefore, it is possible to mark a speech style through a variety of means. A single linguistic feature may be associated with different speech styles in different contexts and the features that mark a speech
style vary in usage by person depending on the context. It would be surprising if an IP-final boundary tone was exclusive to aegyo and universal among speakers. However, it seems plausible that there are boundary tones that are associated with aegyo more than with other speech styles. It seems likely then, that it would be possible to find out which tone or tones occur most often if the tones of a larger number of utterances by a larger number of participants were analyzed using the K-ToBI. A pattern may emerge in a larger study that could not be seen in this study.

Experiment 2

RESULTS

The recordings attained in Experiment 1 were made into clips, each containing one or more utterance. There were 78 clips in total; 26 clips (containing identical or similar utterances) were made from the recordings of three subjects from Experiment 1. These are the same clips that were analyzed to test the hypotheses discussed above. Out of the 78 clips judged there were 64 cases of agreement and 14 cases of disagreement between the judges. The subjects agreed 82.05% of the time in their perception of a clip as “aegyo” or “non-aegyo” (see Table 2). The estimated weighted kappa value for the level of agreement between judges was 0.80. This value indicates that the level of agreement between the subjects was substantially high.

Table 2

Descriptive Statistics for Experiment 2 Judging Task

<table>
<thead>
<tr>
<th>Agreed</th>
<th>Disagreed</th>
<th>Percentage of Agreement</th>
<th>Percentage of Disagreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>64</td>
<td>14</td>
<td>82.05</td>
<td>17.94</td>
</tr>
</tbody>
</table>

Out of 78 total clips listened to by the judges, 29 clips were labeled by one or more of the judges against expectations. Based on the construction of the passages and the contextualized information given via instructions to Experiment 1 subjects before recording, it was expected that clips A, B, N, and O would be rated as “yes” or “unsure,” clips C, D, E, P, Q, and R would be rated as “no” and that the remaining clips would be rated as “yes.” The hypothesis that clips derived from the recordings made in Experiment 1 would be perceived (and thus rated as) aegyo, was met in many of the expected cases, implying that the construction of the Experiment 1 stimuli was at least somewhat successful in eliciting aegyo. It was expected that some clips would not be perceived as aegyo by one judge and might be perceived as aegyo by another judge due to the subjective nature of perception and the varying definitions of aegyo that subjects might have had. The high but not perfect level of concordance between judges was satisfactory, and some instances of ratings that went against expectations were anticipated.
CONCLUSION

As there are few scholarly linguistic studies in English about aegyo to date, it is my hope that this study will prove useful to those interested in aegyo. However, while the results of this study may be informative, they cannot be extrapolated to a larger population. For example, individual differences that might have “averaged” out in a larger study were strong due to the small sample size of this study. Ultimately, this study has remained beneficial in exploring some of the variables that may mark the aegyo speech style.

The fact that the clips were perceived as aegyo or non-aegyo as predicted by both of the Experiment 2 subjects in the majority of cases suggests that the instruments used in Experiment 1 were successful in eliciting the divergent speech styles. Of the three variables quantitatively analyzed and one variable qualitatively examined, there was only one variable, mean pitch, which varied significantly depending on the context. Due to the lack of significant difference between the variables investigated, one hypothesis that came out of this study is that an additional unexplored variable, or a combination of multiple variables, may play an even greater role than prosody in marking aegyo. The variables tested in this study were not an exhaustive list, and other variables, such as duration, nasalization, and palatalization, may also be indicators of aegyo.

Another reason why this study may be limited in its usefulness is its design. The linguistic possibilities that could have been explored through a different research design were curtailed in this study. In Tasks 1 and 4, passages were modeled after the stereotypical contexts that I had been exposed to in Korean media, and therefore were susceptible to my own biases. Also, context information was constructed to attempt to elicit only one type of aegyo, cute and child-like aegyo that is used between two people who are romantically related. However there are a variety of contexts for the use of aegyo, and further research should be conducted giving more attention to other varieties of the speech style in different contexts. It cannot be ignored that the results of this study may not occur in a “natural” environment, outside of the lab. Inevitably, the “hand” of the researcher designing the experiment instruments had some impact on the behavior observed.

Other factors that may have limited this study’s applicability to larger populations were that all of the subjects were current or recently graduated college students, were close in age, and had experience studying one or more foreign languages and had studied abroad. The similarities in performance among the subjects in Experiment 1 and Experiment 2 may not have been found in a subject pool with different occupations, ages, and educational backgrounds. While the applicability of this study to other populations seems small, it still holds great insight for the small population it sampled from. And as Coupland asserts such findings may have been “lost” or “overlooked” in a study of greater size. Further studies must be conducted in the future to help build a more comprehensive understanding of aegyo.

Notes

defined as “남에게 귀엽게 보이는 태도.” I translated “태도” as meaning “manner or attitude”.


6 Ibid. Page 153.


10 The clips from Pairs AN, BO, CP, DQ, and ER were not included in the calculation of the mean absolute slope, pitch range, or mean pitch because they were control clips which were expected to elicit similar if not the same result in both tasks. All other clips, due to the fact that they were “said” by Seong-hwe in Task 1 and contrastively by Young-mi, whose cuteness was emphasized in Task 4, were expected to elicit different results and thus were used for analysis.


12 Some clips contained more than one utterance. For each utterance the IP-final word was divided into syllables. Then the IP-final syllable was coded for. Each utterance and their respective K-ToBI label, rather than each clip, was counted individually for Table 2.

13 For example, another bias in the script was that the dialogue from which contrastive utterances were drawn took place between a girl and her boyfriend in each task. Furthermore, non-gender neutral words such as “oppa” (meaning literally older brother but boyfriend in this context) were used to emphasize the relationship between the girl and her boyfriend. It may be that subjects (unconsciously) attempted to speak in a manner that they thought best resembled their prototypical idea of how a female would act toward a male. Subjects may have also spoken in a manner that they thought best fit how any gendered person would act towards a male. However, it seems unlikely that subjects would have spoken in a way that they thought best fit the behavior of a male who was speaking towards a male. And of course other relations, such as a female speaking with a female, were unlikely to occur due to the construction of the passage and the context clues embedded. Therefore, it is impossible to ascertain if this data on aegyo fits other contexts as well.

14 Coupland writes: “Styling is part of the make-up of public as well as private discourses, and there is no overarching need to restrict sociolinguistics to sampling the speech of ‘ordinary folk’. Whichever speakers we settle on as informants—for a wide variety of reasons—the individual case needs to be addressed as well as the general tendency. This because aggregation rounds down our understanding of stylistic processes. It often blurs the potential for analytic insight. Single-case analyses are more likely to allow an adequate sensitivity to context and contextualization, where we can come to understand what the styling of variation can achieve. There is the possibility of generalizing form single-case analyses, but it involves generalizing to what is stylistically possible, rather than to ‘what people typically do” (27-28).