THE RELATIONSHIP BETWEEN LISTENER ATTITUDES AND THE
COMPREHENSION OF NONNATIVE-ACCENTED SPEECH

by

Julia Vonessen

A Senior Honors Thesis Submitted to the Faculty of
The University of Utah
In Partial Fulfillment of the Requirements for the
Honors Degree in Bachelor of Arts

In

Linguistics

Approved:

________________________________________  ____________________________________________
Rachel Hayes-Harb, PhD                      Scott Jarvis, PhD
Thesis Faculty Supervisor                   Chair, Department of Linguistics

________________________________________  ____________________________________________
Aaron Kaplan, PhD                           Sylvia D. Torti, PhD
Honors Faculty Advisor                      Dean, Honors College

May 2020
Copyright © 2020
All Rights Reserved
The traditional view that the responsibility for the intelligibility of nonnative speech rests solely with the nonnative speaker is challenged by studies demonstrating the contribution of listener factors to nonnative speech intelligibility; these factors include the listener’s language background (Hayes-Harb, Smith, Bradlow & Bent, 2008); exposure to nonnative speech (Clarke & Garrett, 2004), and experience with a particular nonnative accent (Kennedy & Trofimovich, 2008). In addition, native speakers often exhibit negative attitudes toward nonnative speakers (Gluszek & Dovidio, 2010), and Ingvalson, Lansford, Federova & Fernandez (2017) revealed that such attitudes may be associated with a reduced ability to comprehend nonnative speech. Ingvalson et al demonstrated that native English speakers with more negative attitudes toward specific nonnative talkers exhibit a reduced ability to understand those talkers. However, their study involved only one talker for each native language, conflating talker and native language. We conducted an adaptation of Ingvalson et al.’s work, including eight talkers for each of three native languages (Korean, Mandarin, and Spanish).

Native English speakers (n=60) participated in attitude judgment and sentence transcription tasks. Each participant heard a speech sample from half of the 24 nonnative talkers (3 native languages * 8 talkers/language) and rated the speech for ten talker and speech characteristics, and then transcribed 60 sentences produced by the same talkers (5 sentences/talker); these were embedded in 6-talker babble (Van Engen & Bradlow, 2007). The talker and speech ratings were highly correlated and were combined into a single attitude score. A hierarchical linear regression showed that attitude accounted for an additional 3.7% of the variance in transcription accuracy after listener age, gender,
bilingual status, and knowledge of any L2s was accounted for. The strength of this relationship was greater for Spanish than for Mandarin, and attitude did not contribute significantly to the variance in intelligibility for Korean. These differences in the relationship between attitudes and intelligibility provides evidence for the role of attitudes toward accents in addition to attitudes towards talkers.
# TABLE OF CONTENTS

- ABSTRACT ii
- INTRODUCTION 1
- METHODS 3
- RESULTS 9
- DISCUSSION 15
- REFERENCES 19
- APPENDIX 22
INTRODUCTION

The presence of a nonnative accent is one of many variables that impact speech intelligibility, which Munro and Derwing (1995) define as “the extent to which a speaker’s message is actually understood by a listener” (Munro & Derwing, 1995, p. 289). Intelligibility is often examined in the context of factors relating to the speaker and the actual speech sample, such as speech produced by individuals with dysarthria (e.g., Patel et al, 2014) and the presence of noise in the speech sample (Munro, 1998). Speech produced by nonnative-accented talkers has also been examined and tends to be less intelligible to native speakers than is speech produced by native-accented talkers (Munro, 1998). However, research has demonstrated that listeners vary in their ability to understand nonnative-accented speech, depending on experience with nonnative accents (e.g., Clarke & Garrett, 2004; Kennedy & Trofimovich, 2008) and other factors (see, e.g., Baese-Berk et al., 2013; Bradlow and Bent, 2008; Clopper and Pisoni, 2004). This variability in listener ability to perceive nonnative-accented speech argues the importance of listener factors on intelligibility, in addition to talker factors.

One listener-related factor which could contribute to the intelligibility of nonnative-accented speech is listeners’ attitudes towards nonnative talkers and their speech, as negative attitudes towards nonnative talkers are widespread (e.g., Gluscek & Dovidio, 2011). Listeners have been shown to “prefer” the accents of native speakers over those of non-native speakers (Pantos & Perkins, 2013) and, among non-native accents, “prefer” those which are generally more intelligible (Breshanan et al., 2002). In addition, the presence of accents appears to be crucially intertwined with listener
attitudes: accentedness is associated with a decrease in credibility (Lev-Ari & Keysar, 2010), and listener attitudes affect native speakers’ perceptions of the success of interactions with nonnative speakers (Lindemann, 2000). Listener attitudes towards certain ethnic groups have also been shown to affect intelligibility; Babel and Russell (2015) found that native English listeners demonstrated reduced comprehension of the very same speakers when their speech samples were visually associated with images of Chinese Canadians as opposed to images of white Canadians.

The influence of listeners’ explicit attitudes toward accented talkers on speech intelligibility has also been investigated. Ingvalson, Lansford, Federova, and Fernandez (2017) investigated the existence of a relationship between listener attitudes toward specific nonnative talkers and the intelligibility of those talkers’ speech. Native English speakers listened to recordings from five nonnative-accented talkers. They rated the talkers and their speech on several different characteristics and transcribed phrases produced by those talkers. The attitude ratings accounted for a small but significant percentage of the variance in transcription accuracy rates, even after listener age, hearing acuity, and cognitive performance were accounted for. Ingvalson et al concluded from their findings that listeners’ speech perception was indeed related to their attitudes towards individual talkers (Ingvalson et al., 2017). However, the tasks only used one talker for each of five native languages, thus conflating talker and the talker’s native language.

The present study follows up on Ingvalson et al’s work by including 8 talkers for each of 3 native languages, in order to be able to distinguish between talker and native language background. It also examines the effect of several listener factors on the
existence of the relationship between listener attitudes and intelligibility. The use of
different materials also allows us to test the robustness of the effect that they found.

Research Questions

The research question that we investigate here is whether there is a relationship
between listeners’ comprehension of speech produced by nonnative talkers and their
attitudes toward those talkers and their speech. We also investigate whether that
relationship is affected by the native language of the talker, the listener’s bilingual status,
whether or not the listener speaks any second languages, and the listener’s gender.

1. Is there a relationship between listeners’ attitudes towards nonnative talkers and
   their speech and how well they comprehend speech produced by those talkers?
2. Does the native language of the talker affect this relationship?
3. Do the following listener characteristics affect this relationship?
   a. Whether or not the listener is bilingual
   b. Whether or not the listener speaks any second languages (L2s)
   c. The listener’s gender

METHODS

Speech Materials

24 nonnative English speakers (ages 20-43, mean age 26.4) were selected from a
previously existing database of recordings of nonnative English speakers recruited from
the University of Utah community. Eight talkers were randomly selected for each of three
languages: Korean, Mandarin Chinese, and Spanish. Talkers who did not have a
discernible nonnative accent or who had a speech, language, hearing, or other
neurological disorder were not considered. 16 of the talkers were female and 8 were male. All of the selected talkers studied English in high school; 15 out of the 24 studied English in college. Nineteen talkers reported having lived in an English-speaking country, with length of stay ranging from two months to 23 years and a mean length of stay of 5.6 years. Only one of the talkers reported having no informal out-of-classroom exposure to English. No attempt was made to control for the proficiency levels of the talkers. Throughout this paper, we will use “talker” to refer specifically to these 24 participants who produced the recordings used in the experiment.

Each of the talkers produced three readings of the Please Call Stella Passage (Weinberger & Kunath, 2011). For each talker, the second out of the three productions was used. Each of the talkers also produced 21 lists of 19 Bamford-Kowal-Bench (BKB) sentences (Bench et al., 1979). The BKB sentences were very short; each contained 2-4 content keywords. The Please Call Stella script and an example BKB production list are presented in the Appendix.

In order to select 120 BKB sentences for use in the experiment, each of the 21 production lists was divided into three blocks of five sentences, with the two first sentences and two last sentences in each list regarded as buffer sentences and not used. 24 of the resultant 63 blocks were randomly selected. For the first ordering, each of these blocks was randomly assigned one to each talker, so that five sentence productions were used from each talker. A second ordering was constructed as well, with the same 24 blocks assigned to a different talker, so that a randomly selected half of the blocks that were assigned to a native Spanish speaker in the first ordering were assigned to a randomly selected native Mandarin speaker in the second ordering, while the other half
were assigned to a randomly selected native Korean speaker, and so forth. Each sentence was therefore produced twice, once per ordering and each time by a speaker with a different native language background.

Some of the sentences assigned to a particular talker had errors in production which made them unsuitable for use in the study (n=23). This usually resulted from a keyword being produced incorrectly (e.g. “police officer” instead of “policeman”). Replacement sentences were selected from randomly chosen sentence blocks so that 120 sentences were still present for each of the two orderings, 10 productions total from each of the 24 talkers.

All of the sound files were normalized to 70 dB. Pilot data collected from 109 native English participants showed a ceiling effect on transcription accuracy for the BKB sentences. In order to make the task more difficult, the BKB sentences were embedded in 6-talker babble with a signal-to-noise ratio of 0 (Van Engen & Bradlow, 2007).

Participants

Sixty native English speakers were recruited from University of Utah linguistics courses and received course credit for participation. They were 18-62 years of age with a mean age of 23.1, and they received course credit for participation. Thirty-one of the listeners were female and twenty-nine were male. In addition, five identified a native language in addition to English, what we will refer to as bilingual, and forty-four identified experience with at least one L2.

None of the participants spoke any of the target languages with parents or childhood caregivers. A listener’s responses to a particular talker were excluded if the
listener had L2 experience with that talker’s L1. We excluded those responses from one listener for Korean, five for Mandarin, and twenty-five for Spanish.

**Procedures**

The survey was conducted online using the Qualtrics survey platform. Participants were asked to use headphones and to complete the study in a quiet environment. To ensure that those conditions were at least partly met, participants first completed a soundcheck where they had to identify four English words at 100% accuracy. Once they passed the soundcheck, the participants completed a consent form. They completed two tasks: an attitude rating task and a transcription task. For both tasks, each listener was randomly assigned to the same 12 talkers, four per language.

For the attitude rating task, listeners heard the “Please call Stella” passage from each of their 12 talkers, in random order. They rated the speech and the talker on 10 different characteristics on a 5-point Likert scale. This was following the procedure of Ingvalson et al, adapted from Pantos and Perkins (2013). The speech characteristics were pleasantness, romantic qualities, warmth, refinement, and ease of understanding. They were rated from LEAST/LOWEST to MOST/HIGHEST, where a rating of LEAST/LOWEST corresponded to a numerical value of 5, and a rating of MOST/HIGHEST corresponded to a numerical value of 1. Figure 1 shows the screen that participants viewed during this task. The ratings for these five speech characteristics were averaged to produce a speech attitude score. A lower score corresponded to a more positive rating. The talker characteristics, education, class status, pleasantness, friendliness, and confidence, were rated on the same scale. The ratings for the five talker
characteristics were averaged to produce a talker attitude score. Again, a lower score corresponded to a more positive rating.

For the transcription task, listeners were randomly assigned to one of two sentence lists. Using a computer keyboard, they transcribed 60 sentences, five from each talker. The sentences were blocked by talker and presented in random order. Figure 2 shows the screen that participants saw during this task.

Transcription Coding

The first author coded all 3600 responses according to the following coding scheme:

1. Each of the BKB sentences contained 2-4 keywords; only these keywords were taken into account. Intelligibility scores were comprised of mean percent keywords correct.

2. Keywords that were split into two words (“rain coat” for “raincoat”) or two keywords that were combined into one word (“matchboxes” for “match boxes”) were counted as correct.

3. Numbers written with numerals instead of spelled out (“9” instead of “nine”) were counted as correct.

4. Homophones and common alternate spellings were counted as correct (“road” for “rode” and “paid” for “payed” were fine, but “wrote” for “rode” was not).

5. Contractions were counted as correct as long as no keyword meaning was lost. For example, “the car engine’s running” for “the car engine is running” was counted as correct, but “the car engines running” was not.
On the basis of the recording you hear, please rate the speaker’s speech as follows:

<table>
<thead>
<tr>
<th>MOST/HIGHEST rating</th>
<th>LEAST/LOWEST rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pleasantness</td>
<td></td>
</tr>
<tr>
<td>Romantic qualities</td>
<td></td>
</tr>
<tr>
<td>Warmth</td>
<td></td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
</tr>
<tr>
<td>Ease of understanding</td>
<td></td>
</tr>
</tbody>
</table>

On the basis of the recording you hear, please rate the speaker as follows:

<table>
<thead>
<tr>
<th>MOST/HIGHEST rating</th>
<th>LEAST/LOWEST rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education</td>
<td></td>
</tr>
<tr>
<td>Class status</td>
<td></td>
</tr>
<tr>
<td>Pleasantness</td>
<td></td>
</tr>
<tr>
<td>Friendliness</td>
<td></td>
</tr>
<tr>
<td>Confidence</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1 – Participant view of attitude rating task

Transcribe the sentence that you hear.

Figure 2 – Participant view of sentence transcription task
6. If words did not have the correct inflection ("talk" instead of "talked"), they were marked as incorrect.

7. If two options were put down for the keyword ("(slipped)/(slept)" for "slipped"), the word was marked as incorrect. Likewise, if a question mark was used to indicate uncertainty ("jar (???)" for "jar"), the preceding word was marked incorrect.

8. Anything that was not a properly-spelled English word was counted as incorrect.

For the purpose of determining the reliability of the coding scheme, a second rater (the second author) coded a random selection of 10% of responses. Intraclass Correlation (ICC) estimates and their 95% confident intervals were calculated using SPSS statistical package version 25 (SPSS Inc, Chicago, IL) based on a mean-rating (k = 2), absolute-agreement, 2-way random-effects model. The average measures ICC was .997, which is considered “excellent reliability” (Koo & Li, 2016); F(359,359)=681.274, p<.0005, CI[.996, .998]. On the items where the rater scores diverged (n=3), the first author’s score was used for further analysis.

RESULTS

The talker attitude scores and speech attitude scores were highly correlated (Pearson r=.777, p<.0005, n=596). Figure 3 shows a scatterplot of these scores with speech attitude scores on the vertical axis and talker attitude scores on the horizontal axis. Like Ingvalson et al (2017), we averaged these scores into a single composite attitude score and used that attitude score for subsequent analyses.
The distribution of the proportion of keywords correctly transcribed for each talker by each listener was normal, as shown in the histogram in Figure 4. Embedding the sound files in 6-talker babble successfully lowered the ceiling effect found in the pilot data.
Figure 5 presents a scatterplot with the composite attitude scores on the vertical axis and transcription accuracy on the horizontal axis. A hierarchical linear regression showed that where the listener’s age, gender, bilingual status, and knowledge of any L2s did not account for any of the variance in transcription accuracy ($F(4,591)=0.705$, $R^2=.005$, $p=.589$), the addition of composite attitude scores to the model resulted in the model accounting for a significant percentage of the variance in transcription accuracy ($F(5,590)=5.157$, $\Delta R^2=.037$, $p<.0005$). Table 1 presents the results of this hierarchical linear regression. Attitude was the only variable that was a significant contributor to either model ($b=-.048$, $SE=.010$, $p<.0005$). The fact that the regression coefficient for attitude is negative is expected, because a numerically lower attitude score is a more positive rating, providing evidence for more positive attitudes being related to greater intelligibility.

![Figure 5 – Scatterplot of composite attitude scores and transcription accuracy](image-url)
Table 1: Results of hierarchical linear regression for all listeners

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>b</th>
<th>SE</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1: Listener characteristics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>.002</td>
<td>.064</td>
<td>.121</td>
</tr>
<tr>
<td>Gender (1=F, 2=M)</td>
<td>.004</td>
<td>.010</td>
<td>.817</td>
</tr>
<tr>
<td>Bilingual (1=yes, 0=no)</td>
<td>-.003</td>
<td>-.004</td>
<td>.924</td>
</tr>
<tr>
<td>L2s (1=yes, 0=no)</td>
<td>-.008</td>
<td>-.018</td>
<td>.657</td>
</tr>
<tr>
<td>Model 2: Attitudes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>.002</td>
<td>.002</td>
<td>.168</td>
</tr>
<tr>
<td>Gender (1=F, 2=M)</td>
<td>.024</td>
<td>.017</td>
<td>.159</td>
</tr>
<tr>
<td>Bilingual (1=yes, 0=no)</td>
<td>-.006</td>
<td>.030</td>
<td>.829</td>
</tr>
<tr>
<td>L2s (1=yes, 0=no)</td>
<td>-.019</td>
<td>.017</td>
<td>.276</td>
</tr>
<tr>
<td>Attitude</td>
<td>-.048</td>
<td>.010</td>
<td>.000</td>
</tr>
</tbody>
</table>

Table 2 presents the mean attitude and intelligibility scores for all listeners as well as the mean scores for listeners split by gender, L2 status, and bilingual status. It also presents the mean attitude and intelligibility scores for each of the talker languages. T-tests for equality of means with equal variances not assumed showed that females and males did not differ significantly in transcription accuracy (p=.680), but that males gave significantly harsher ratings (p<.0005). Likewise, listeners without L2s rated the talkers more harshly than listeners with L2s (p<.0005), but the transcription accuracy between the two groups did not differ (p=.655).

We did not perform t-tests for differences in the bilingual status groups, because there were only 5 bilingual listeners. However, t-tests for equality of means showed that Spanish was rated significantly higher than both Korean (p<.0005) and Mandarin.
Table 2: Mean attitude and intelligibility scores

<table>
<thead>
<tr>
<th></th>
<th>Attitude</th>
<th></th>
<th>Intelligibility</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>All listeners</td>
<td>2.9534</td>
<td>.83949</td>
<td>.6157</td>
<td>.00820</td>
</tr>
<tr>
<td>Listener gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>2.7600</td>
<td>.86127</td>
<td>.6126</td>
<td>.20391</td>
</tr>
<tr>
<td>Male</td>
<td>3.1775</td>
<td>.75525</td>
<td>.6194</td>
<td>.19615</td>
</tr>
<tr>
<td>Listener L2 status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Has L2s</td>
<td>2.8700</td>
<td>.85894</td>
<td>.6132</td>
<td>.19664</td>
</tr>
<tr>
<td>Does not have L2s</td>
<td>3.1286</td>
<td>.77017</td>
<td>.6212</td>
<td>.20795</td>
</tr>
<tr>
<td>Listener bilingual status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monolingual</td>
<td>2.9527</td>
<td>.85247</td>
<td>.6162</td>
<td>.20042</td>
</tr>
<tr>
<td>Bilingual</td>
<td>2.9604</td>
<td>.68128</td>
<td>.6104</td>
<td>.19993</td>
</tr>
<tr>
<td>Talker language</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Korean</td>
<td>3.0695</td>
<td>.84682</td>
<td>.6320</td>
<td>.19177</td>
</tr>
<tr>
<td>Mandarin</td>
<td>2.9618</td>
<td>.80415</td>
<td>.5828</td>
<td>.19648</td>
</tr>
<tr>
<td>Spanish</td>
<td>2.7443</td>
<td>.84762</td>
<td>.6401</td>
<td>.21407</td>
</tr>
</tbody>
</table>

(p=.016), but the ratings for Korean compared to Mandarin did not differ significantly (p=.164). Transcription accuracy for the Mandarin talkers was significantly lower than for the Korean (p=.007) and Spanish (p=.011) talkers, while the transcription accuracy for the Korean and Spanish talkers did not differ significantly (p=.712). However, it is important to keep in mind that no attempt was made to control for proficiency of the talkers.

Three hierarchical linear regressions were performed separately for the attitude and transcription accuracy scores for each talker language. These showed that for Korean,
attitude did not contribute significantly to the variance in transcription accuracy
\( (F(5,230)=0.983, \Delta R^2=.011, p=.107) \). For Mandarin, however, attitude accounted for
1.8\% of the variance in transcription accuracy after listener gender, age, L2 status, and
bilingual status were accounted for \( (F(5,214)=1.547, \Delta R^2=.018, p=.045) \). For Spanish,
attitude accounted for an additional 14.8\% of the variance in transcription accuracy
\( (F(5,134)=5.062, \Delta R^2=.148, p<.0005) \).

Separate hierarchical linear regressions were not performed for bilingual and
monolingual listeners because there were only 5 bilingual listeners. Hierarchical linear
regressions were performed separately for listeners with and without L2s. For listeners
without L2s, attitudes accounted for an additional 3.1\% of the variance in transcription
accuracy \( (F(4,187)=3.354, \Delta R^2=.031, p=.014) \). Both attitudes \( (b=-.048, SE=.019, p=.014) \)
and listener gender \( (b=.067, SE=.032, p=.037) \) contributed significantly to the model. For
listeners with L2s, attitudes accounted for an additional 3.6\% of the variance in
transcription accuracy \( (F(4,399)=4.759, \Delta R^2=.036, p<.0005) \), but attitude was the only
significant predictor \( (b=-.045, SE=.012, p<.0005) \).

Hierarchical linear regressions were also performed separately for female and
male listeners. For female listeners, attitude accounted for an additional 1.7\% of the
variance in transcription accuracy \( (F(4,315)=4.193, \Delta R^2=.017, p=.019) \). Both attitudes
\( (b=-.032, SE=.013, p=.019) \) and listener bilingual status \( (b=-.155, SE=.052, p=.003) \)
contributed significantly to the model. For male listeners, attitude accounted for an
additional 7.8\% of the variance in transcription accuracy \( (F(4,271)=9.274, \Delta R^2=.078, \)
p<.0005). Attitudes \( (b=-.073, SE=.015, p<.000) \), listener L2 status \( (b=-.059, SE=.024, \)
p=.013), and listener age ($b=.004, SE=.002, p=.016$) contributed significantly to the model.

DISCUSSION

This study investigated the existence of a relationship between intelligibility and listener attitudes towards nonnative talkers and their speech. We found evidence for the existence of such a relationship, and found that the relationship goes in the expected direction, with more positive attitudes corresponding to increased intelligibility. This replicates the findings of Ingvalson et al (2017) and fits with other research that provides evidence for listener factors on intelligibility, such as experience with L2 speech (Kennedy & Trofimovich, 2008) and the listener’s native language (Hayes-Harb et al, 2008).

We also investigated the effect of talker language on the relationship between attitudes and intelligibility and found that the relationship was stronger for the Spanish talkers than for the Korean and Mandarin talkers. Given the status of Spanish in the U.S. and the greater number of native Spanish speakers than native Mandarin or Korean speakers, this is a finding that is worth following up on. Spanish talkers were also rated more positively than the talkers of the other two languages. This aligns with previous research that has shown that native English speakers rate Spanish-accented talkers more positively than talkers whose L1 is from an Asian language family (e.g., Sales, 2012).

The last research question focused on the effect of several listener characteristics on the relationship between attitudes and intelligibility. Due to the small number of bilinguals, we were unable to draw a conclusion on whether the listener’s bilingual status
impacts the relationship between listener attitudes and intelligibility. There was no
difference in the strength of the relationship between attitudes and intelligibility for
listeners with and without L2s, despite the fact that listeners without L2s rated the talkers
more harshly than the listeners with L2s. Under the assumption that individuals who have
acquired L2s have in general more experience with nonnative-accented speech than those
who have not, this last finding aligns with research that shows that listeners with
experience with nonnative-accented speech rate nonnative-accented talkers more
positively on characteristics such as comprehensibility (e.g., Sales, 2012).

As the relationship between attitudes and intelligibility was stronger for males
than for females, we did find that listener gender had an impact on the relationship
between attitude and intelligibility. We also found that males rated the talkers more
harshly than females did. These findings fit with various studies that have shown that
females and males respond differently to the same language stimuli in various settings
(Podberesky et al, 1990; Brown & Cichocki, 1995; O’Loughlin, 2002; Grondelaers et al,
2010).

In contrast to Ingvalson et al (2017), we did not find that age was a significant
predictor of intelligibility. However, half of Ingvalson et al’s participants were older
adults, while our study had only one participant above the age of 35; this may play a role
in that finding.

In summary, our study provided evidence that more positive attitudes toward
nonnative talkers and their speech are associated with increased intelligibility of speech
produced by those talkers. In addition, the results suggested that the strength of that
relationship differs by native language, and that the strength of the relationship appears to
increase for male listeners and listeners with experience with L2s. Due to low sample size, we were unable to draw any conclusion about the listener’s bilingual status.

Although we set out to tease apart attitudes towards talkers and attitudes toward accents, this single experiment does not accomplish this, because the same talkers are present in both the attitude rating and transcription tasks, thus preserving the confound of talker and accent. Data collection is complete for a second experiment with the same methodology, except that the listeners heard different talkers in the transcription task and the attitude rating task, thus removing the confound in question. If this experiment shows a differing relationship between attitude and intelligibility for the three talker native languages, that would provide evidence for attitudes toward a particular native language background driving the relationship between attitudes and intelligibility. If not, that would provide evidence for attitudes towards particular talkers driving the relationship between attitudes and intelligibility, supporting Ingvalson et al.’s (2017) finding that listener attitudes toward particular talkers reflects attitudes toward talkers and not attitudes toward accents.

In their model, Ingvalson et al (2017) accounted for listener age, hearing acuity, and cognitive factors, whereas our model only accounted for age. A fuller replication, which would include measures of these factors, is planned; however, data collection has been suspended due to the current health crisis.

This study provided further evidence that the listener contributes to intelligibility, supporting the idea that intelligibility is the responsibility of both the speaker and the listener. This is excellent news, as it suggests the possibility of multiple paths to
enhancing the intelligibility of nonnative-accented speech – from both sides of the interaction.
REFERENCES


APPENDIX

This appendix presents the production materials used to obtain recordings from the nonnative talkers: the Please Call Stella Passage (Weinberger & Kunath, 2011) and an example production list of BKB sentences (Bench et al, 1979).

1. Please call Stella. Ask her to bring these things with her from the store: Six spoons of fresh snow peas, five thick slabs of blue cheese, and maybe a snack for her brother Bob. We also need a small plastic snake and a big toy frog for the kids. She can scoop these things into three red bags, and we will go meet her Wednesday at the train station.

2. Please call Stella. Ask her to bring these things with her from the store: Six spoons of fresh snow peas, five thick slabs of blue cheese, and maybe a snack for her brother Bob. We also need a small plastic snake and a big toy frog for the kids. She can scoop these things into three red bags, and we will go meet her Wednesday at the train station.

3. Please call Stella. Ask her to bring these things with her from the store: Six spoons of fresh snow peas, five thick slabs of blue cheese, and maybe a snack for her brother Bob. We also need a small plastic snake and a big toy frog for the kids. She can scoop these things into three red bags, and we will go meet her Wednesday at the train station.
List 11

1. Many horses like hay.
2. The candy shop was empty.
3. The dogs went for a walk.
4. She is washing her dress.
5. The lady stayed for lunch.
6. The driver waited by the corner.
7. They finished the dinner.
8. The policeman knows the way.
9. The little girl was happy.
10. He wore his yellow shirt.
11. They are coming for Christmas.
12. The cow gave some milk.
13. The boy got into bed.
14. The two farmers are talking.
15. Mother picked some flowers.
16. A fish swam in the pond.
17. Father wrote a letter.
18. The mailman delivered a box.
19. The woman spoke loudly
Name of Candidate: Julia Vonessen

Birth date: January 30, 1998

Birth place: Missoula, Montana

Address: 109 Terrace Way
Missoula, MT 59803