The Effects of Accent Familiarity and Language Attitudes on Perceived English Proficiency and Accentedness

Jocelyn Carnicle & Becky Huang

University of Texas at San Antonio, San Antonio, Texas, United States of America

ABSTRACT

If accents are a permanent indicator of difference, to what degree do our accents determine how we are perceived by others? The goal of this project was to examine the relationship between raters' accent familiarity, attitudes toward non-native accents, and their judgments of non-native speech. This project was designed to extend the Huang, Alegre, & Eisenberg (2016) study with several methodological improvements. The study included five groups of raters who vary in their familiarity with Korean and Arabic accents (No Familiarity, Korean Heritage Raters, Korean Non-Heritage Raters, Arabic Heritage Raters, Arabic Non-Heritage Raters). There were 10 raters in each group (n = 50), and all participants consisted of undergraduate or graduate students who were born in the U.S. or immigrated to the U.S. before the age of twelve. Participants listened to 24 speech samples selected from Educational Testing Service's TOEFL iBT public database, and rated each sample on magnitudes of perceived proficiency and accentedness. After rating the speech samples, participants completed a survey on their demographic background information, attitudes toward accents, rating tendency, and beliefs/perceived cultural factors. Finally, two participants from each group were randomly selected to participate in a face-to-face think-aloud follow-up study to discuss their cognitive processes during rating. Because foreign accents can be subject to negative perceptions and linguistic profiling, understanding the potential sources of biases is critical for removing such biases and improving human communication and interactions.

KEY WORDS:

INTRODUCTION

Accents are detected as audible difference to listeners and often serve as an index of a speaker's native or non-native status, which can be linked to many resentment ideologies and even discriminatory linguistic profiling (Makoni, 2003). Understanding perceptions of foreign accents is then an imperative subject of research in employment, overcoming communication obstacles between monolingual and bilingual interlocutors, and the instruction and assessment of English as a second/foreign language (ESL/EFL).

In particular, ESL/EFL speaking assessments utilize human judgments in determining test takers' language proficiency. Rater effect, defined as construct-irrelevant variations associated with rater background and characteristics, thus plays a critical role in the reliability and validity of the speaking assessment. While many studies have focused on the effects of accent familiarity on rater perceptions, research on the source of accent familiarity such as heritage or familial ties, is much more limited. In a seminal study, Winke & Gass (2013) found that heritage learners, i.e. someone who has cultural or ethnic ties to a second language and family members who are native speakers of that heritage language, rated speakers with the corresponding first language (L1) highest. The authors speculated that, when raters are more familiar with the test taker's L1, they tend to orient to the learners' talk in a biased way, which compromises test reliability by promoting leniency in scoring. The authors called for further research to specifically examine if heritage learners are more prone to bias than non-heritage learners.

In a recent cross-linguistic study on the role of accent familiarity in raters' evaluations of non-native speech, Huang, Alegre, & Eisenberg (2016) investigated whether familiarity with a particular accent, in this case Spanish and Chinese accents, and type of familiarity (heritage vs. non-heritage) motivated the way participants rated non-native speech samples. Huang et al found that familiarity with an accent facilitated identification of that particular accent; however, the study found no significant differences in the way the source of the participants' familiarity influenced how they rated accented speech samples. Participants in this study did, however, self-report feeling more sympathetic in the way they rated accents with which they were familiar. This discrepancy between raters' perception of their own bias and their assigned ratings is especially intriguing. The mixed results could be attributed to the context of the study: as the participants in the study were recruited in San Antonio where Spanish accent is prevalent, the Chinese raters in the study have some familiarity with Spanish accents.

Given the limited research on the effect of the source or type of accent familiarity on speech evaluation, and the mixed results from the most recent cross-linguistic rater effect study by Huang et al (2016), the current project is designed to address this topic and fill the gap in the existing literature. The current project has two goals: to test the results of Huang et al. (2016) study with improved methodology, and to systematically investigate the relationship between accent familiarity, language attitudes, and rating behavior. The three methodological improvements are: 1) the current study will utilize a fully crossed design and add an additional baseline group with no familiarity; 2) the study will use Arabic and Korean, two very distinct language groups that are of similar levels of status and prevalence in the area, to avoid confounding variables such as the geography where data are collected and the dominant accents in the location; 3) this study will collect information about raters' beliefs about and attitudes toward non-native speakers to examine their effect on raters' judgments and rating processes.

RESEARCH QUESTIONS

The research questions that guided this study are as follows:

Research question 1: Does raters' accent familiarity relate to their accuracy in identifying non-native accents?

Research question 2: Do raters' accent familiarity and source of accent familiarity relate to the reliability of their rating?

Research question 3: Do raters' accent familiarity and source of accent familiarity relate to the degree of severity of their rating?

Research question 4: Do raters' beliefs about and attitudes toward non-native accents relate to the severity of their ratings?

MATERIALS AND METHODS

Speech Files

The speech samples rated in the online study were selected from Educational Testing Service's (ETS) Test of English as a Foreign Language (TOEFL©) iBT public database. Ten speech samples from Korean language background and ten from Arabic language background were selected for the rating study. Samples from each language background were matched on gender and official TOEFL score provided by ETS (Range = 1-4). An additional four speech samples from a variety of other native language backgrounds were included to diversify the variety of foreign accents in the speech sample. All samples consisted of a spontaneous, narrative response to a standard prompt, and the length of each sample was approximately 45 seconds.

Participants

Rater participants included undergraduate or graduate students between the ages of 18 and 35, with no hearing or speech disorders. To ensure advanced English language proficiency, all rater participants were required to either be have been born in the United States or have immigrated to the United States before the age of twelve. Participants fell in one of five groups of raters that vary in their familiarity with Korean and Arabic accents: 1) No Familiarity Comparison Group (No familiarity with Korean or Arabic accents), 2) Korean Heritage Rater Group (Familiarity with Korean accents through heritage ties), 3) Korean Non-Heritage Rater Group (Familiarity with Korean accents through formal language instruction), 4) Arabic Heritage Rater Group (Familiarity with Arabic accents through heritage ties), and 5) Arabic Non-Heritage learner (Familiarity with Arabic accents through formal language instruction). There will be 10 raters in each group (total n = 50), and gender will be balanced within each group. Raters were recruited from numerous sources, including Arabic and Korean classes at UTSA, flyers on the university campuses, the UTSA Honors College social media pages, and through word of mouth, or snowball sampling. A subsample were randomly selected (n = 2 from each group) to participate in a follow-up, think-aloud study.

Instrument

Each rater evaluated all 20 speech samples online along two dimensions, "Overall English Proficiency" and "Foreign Accents," using a 1-7 Likert scale. Each scale had three descriptors in the midpoint (4) and on both endpoints (1, 7). The descriptors were "low English proficiency" (1), "intermediate English proficiency" (4), and "high English proficiency" (7) for the "Overall English Proficiency" dimension, and "strong foreign accent (1), "some foreign accent" (4), and "native English speaker" for the "Foreign Accents" dimension.

Upon completion of the rating study, each participant completed a background survey. The background survey included questions about the participants' language learning history, and their familiarity with and exposure to non-native accents. Participants were also asked to describe the features, such as pronunciation or word choice, that they attended to most when rating the speech samples on overall proficiency and foreign accent dimensions. The survey also included questions about participants' attitudes toward non-native accents and beliefs/perceived cultural factors (Hsu, 2015; Meerleer, 2012).

Procedure

Participants were told they were participating in a study that examines the effect of raters' accent familiarity on their judgments of non-native speakers' language proficiency. After giving their consent, all participants completed the rating study online in a single 1.5-hour session. Participants first listened and rated the 24 speech samples, one at a time, on Overall Proficiency using a 1-7 Likert scale. Then participants proceeded to rate these same files, again one at a time, on the Foreign Accent

dimension using a 1-7 Likert scale. Participants then filled out a background survey at the end of the study. They received two course credits or a $10 Amazon gift card for their participation. Ten participants from each of the five groups were randomly selected for the follow-up think-aloud study portion. These selected participants met with a researcher in person within 72 hours of finishing their online rating sessions. The think-aloud study provided information about the cognitive processes raters experienced during ratings. The participants who participated in the think-aloud portion received an additional $10 cash for their time.

EXPECTED RESULTS

We expect raters' accent familiarity to facilitate their identification of the specific non-native accents they are familiar with. In particular, for Arabic accents, the Arabic Heritage raters will perform better than the Arabic Non-Heritage raters in the identification task, and the Arabic Non-Heritage raters will perform better than the other three rater groups. We expect the same patterns to occur for the identification of Korean accents (Korean Heritage > Korean Non-Heritage > Arabic Heritage & Non-Heritage raters > No Familiarity). We also expect the No Familiarity group to perform the lowest among the five groups in the identification task.
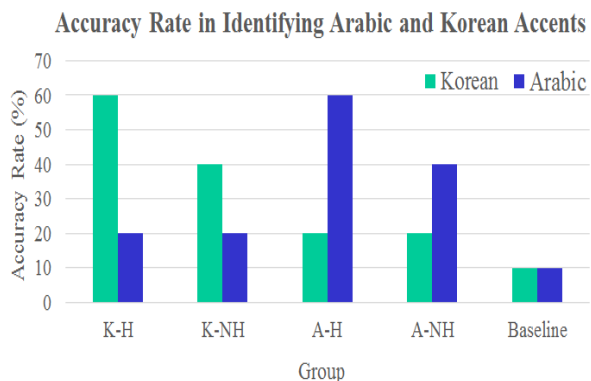

Figure 1. Expected Results: Accuracy Rate in Identifying Arabic and Korean Accents

Given that all rater groups are untrained raters, we expect them to display similar degrees of reliability in their rating.

We also predict that raters who are familiar with a specific accent will evaluate that accent more favorably than raters without such familiarity. Furthermore, given that heritage raters have stronger ties with speakers of the accent they are familiar with, we expect heritage raters to assign higher ratings to the non-native accents than non-heritage raters. Heritage raters will also be more lenient toward non-native accents in general than non-heritage raters. Thus for Korean accents, we expect Korean Heritage raters to assign higher ratings than Korean Non-Heritage raters (Korean Heritage > Korean Non-Heritage > Arabic Heritage > Arabic Non-Heritage > No Familiarity). Similar patterns are expected for Arabic accents: Arabic Heritage > Arabic Non-Heritage > Korean Heritage > Korean Non-Heritage > No Familiarity.
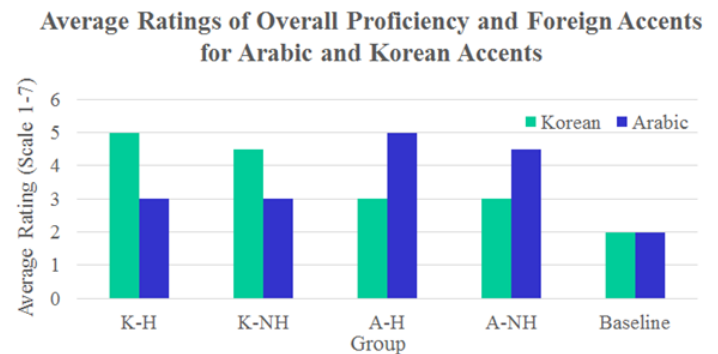

Figure 2. Expected Results: Average Ratings of Overall Proficiency and Foreign Accents for Arabic and Korean Accents

Lastly, we predict that raters' beliefs about and attitudes toward non-native accents, as indexed by their responses to the related survey questions, will correlate with the group categories that they belong to. That is, the heritage rater groups will have more positive attitudes toward non-native accents than the non-heritage rater groups. The No Familiarity rater group is hypothesized to have the least favorable attitudes toward non-native accents. Furthermore, we

expect the variations in raters' language beliefs and attitudes to be the moderator between raters' accent familiarity and the severity of their ratings of non-native speakers' English proficiency and foreign accents.
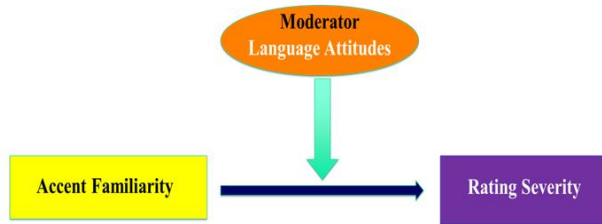

Figure 3. Expected Results: Relationship between Accent Familiarity, Language Attitudes, and Rating Outcomes

SIGNIFICANCE OF THE STUDY

As far as we know, the current study is the first to use a cross-linguistic design that systematically examines the relationship of accent familiarity and language attitudes on rating behavior. Results of the study will contribute to understanding the reliability and validity of using human raters in assessing second language proficiency. Results from the background survey and the think-aloud data will also afford us a better understanding of the effect of raters' attitudes on their judgments as well as their decision-making process. Finally, results from the study will thus help raise awareness of foreign accent biases in the workplace, school campus, and every day communication.

REFERENCES

Beard, B. (2013). No Mere "Matter of Choice": The Harm of Accent Preferences and English-Only Rules. *Texas Law Review, 91*(6), 1495.

Bell, S. M., McCallum, R. S., Kirk, E. R., Brown, K. S., Fuller, E. J., & Scott, K. W. (2009). Psychometric properties of the foreign language attitudes and perceptions survey for college students. *Assessment for Effective Intervention, 35*(1), 54-60.

Cargile, A. C., Giles, H., Ryan, E. B., & Bradac, J. J. (1994). Language attitudes as a social process: A conceptual model and new directions. *Language and Communication, 14*(3), 211-236.

De Meerleer, M. (2012). *Beliefs and Attitudes towards English as a Lingua Franca: Native and Non-native Pronunciation*. (Doctoral dissertation, Master dissertation, Ghent University Faculty of Art and Philosophy, Netherlands).

Derwing, T. M., & Munro, M. J. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching, 42*(4), 476-490.

Gill, M. M. (1994). Accent and Stereotypes: Their Effect on Perceptions of Teachers and Lecture Comprehension. *Journal of Applied Communication Research, 22*(4), 348-361.

Hsu, T. H. -. (2015). Removing bias towards World Englishes: The development of a Rater Attitude Instrument using Indian English as a stimulus. *Language Testing,* 1-23. doi: 10.1177/0265532215590694

Huang, B., Alegre. A., & Eisenberg, A. (2016). A Cross-Linguistic Investigation of the Effect of Raters' Accent Familiarity on Speaking Assessment. *Language Assessment Quarterly, 13*(1), 25-41.

Huang, B. H. (2013). The Effects of Accent Familiarity and Language Teaching Experience on Raters' Judgments of Non-Native Speech. *System, 41*(3), 770-785. doi:10.1016/j.system.2013.07.009

Ingram, P. (2009). Are Accents One of the Last Acceptable Areas for Discrimination? *Journal of Extension, 47*(1) 1-5.

Makoni, S. (2003). *Black linguistics: Language, Society, and Politics in Africa and the Americas.* Psychology Press.

Smith, L. E., & Bisazza, J. A. (1982). The Comprehensibility of Three Varieties of English for College Students in Seven Countries. *Language Learning, 32*(2), 259-269. doi:10.1111/j.1467-1770.1982.tb00971.x

Winke, P., & Gass, S. (2013). The Influence of Second Language Experience and Accent Familiarity on Oral Proficiency Rating: A qualitative investigation. *TESOL Quarterly, 47*(4), 762-789. doi:10.1002/tesq.73