

## Accent Assessment: A Preliminary Study of Scaling Validity

*D. Mike McDaniel, PhD*

Professor of Communication Disorders  
Arkansas State University (USA)

*Richard A. Neeley, PhD*

Professor and Chair, Department of Communication Disorders  
Arkansas State University (USA)

*Julie J. Isaacson, RN, MSN, CCRN*

Associate Professor of Nursing  
Arkansas State University (USA)

*G. Daniel Howard, PhD*

Interim Chancellor  
Arkansas State University (USA)

### Abstract

*Direct Magnitude Estimations (DME) and Equal Appearing Interval (EAI) scaling techniques were used to compare listeners' perceptions of the extent of accent from recorded speech samples of international students enrolled in a United States university who spoke English as a second language. Twenty five international students served as speakers by reading the same brief passage for recording purposes. Twenty five American-born students with no formal training or experience with accents or accent reduction rated the extent of the accent on each of the spoken samples using both scaling techniques. Statistical analysis of the listener's perceptions indicated no significant differences between the DME or EAI scaling procedures and a scatterplot comparing the data sets for each technique produced a significant linear relationship between the data for the two techniques. Based on these findings it would appear that either technique could be employed to obtain a valid assessment of the extent of accent in international students. An EAI scaling technique would appear to be the technique of choice because of the ease of administration. However, future study is needed to refine the technique into a clinical procedure for routine use in assessing the extent of perceived accent and gauging the success of accent reduction intervention.*

**Key Words:** Accent Assessment, International Students, DME, EAI, India

Institutions of higher education in the United States recognize the inherent value of recruiting the best and brightest students from around the world (Institute of International Education, 2011). The number of international students enrolled in American universities is increasing. A 2007 report on enrollment showed a record 623,805 international students were enrolled in U.S. colleges and universities (Association of International Educators, 2009). Recent data reported in the *Chronicle of Higher Education* (Fischer, 2009) indicated that for 2008-2009 there were 672,000 foreign students enrolled in American universities. In 2010/11, the number of international students enrolled in American universities had swelled to 723,000 (Marklein, 2011). These international students facilitate the preparation of American students for a globally connected marketplace by sharing their various cul-

tures, perspectives, and ways of thinking. However, the continuing influx of international students into institutions of higher education in the United States poses verbal communication challenges for the institutions and the international students. English as a second language is often spoken by the majority of these international students and their accents can jeopardize effective communication. Even after gaining sufficient English language literacy skills to successfully matriculate in academic programs in U.S. colleges and universities, the intelligibility of conversational speech among international students may be significantly compromised in academic and social contexts by their accents. Accordingly, institutions of higher education should share some portion of the responsibility for assisting foreign students as they pursue opportunities to overcome accent-based communication barriers.



The American Speech-Language-Hearing Association (2009) defined accent as “the unique way that speech is pronounced by a group of people speaking the same language.” Despite the number of individuals presenting themselves for accent reduction, Shah (2007), in the development of a prototype accent assessment tool, reported that none of the 13 protocols reviewed for assessment of foreign accent were evidenced based, norm referenced, or standardized. Accent reduction, sometimes referred to as elocution or accent modification, should begin with a systematic, valid and objective assessment of the accent prior to formal intervention. In short, a suitable instrument is needed that can reliably quantify deviations in the individual's current accent from the target reference accent and from which progress in accent reduction can be gauged. Before an accent assessment instrument can be developed certain assumptions about the perception of accents in listeners need to be explored to ensure that the accent assessment instrument is appropriate. Accent perception appears to encompass a combination of elements based on quality (kind) and quantity (degree). Again, for an accent assessment instrument to be valid it should reflect a methodology grounded in what is scientifically known about the perception of accent from a listener's perspective.

To determine if accent differences are a matter of kind or a matter of magnitude, a theoretical concept was adapted from the field of psychophysical perception. Stevens (1974) reported works in which perceptions were recorded using both scale and direct magnitude estimation methods. He found that when the nature of perceptual comparisons pertained to differences in magnitude or quantity, the perceptions resided in a prosthetic (amount) continuum and were most effectively sampled using a direct magnitude estimation method. According to Stevens (1974), the correlation between direct magnitude estimations and equal appearing interval (EAI) scaling methods for the same stimuli, determined whether listener perceptions were from a prosthetic or a metathetic continuum. Specifically, he indicated that a linear relationship between interval scaling judgments and direct magnitude estimation judgments of the same stimuli would suggest that either technique would be appropriate for judging perceptions on a metathetic continuum. From a practical standpoint, interval scaling judgments are easier to accomplish because they do not require an anchor or standard reference for comparison judgments. Conversely, Stevens (1974) felt that if the relationship between the two sets of judgments were non-

linear then the use of interval scaling was inappropriate and the dimension of the perception was prosthetic. Previous investigations have compared DME and EAI scaling for other aspects of communication such as vowel roughness (Toner and Emanuel, 1989); stuttering (Schiavetti, Sacco, Metz, and Sitler, 1983; Berry and Silverman, 1972); hearing impairment (Schiavetti, Metz, and Sitler, 1981); dysarthria (Weismer and Laures, 2002); and nasality (Zraick and Liss, 2000).

The purpose of this study was to investigate the nature of perceived accent differences among international (viz., Indian) speakers by listeners unfamiliar with the nature of accents and accent reduction. It was hoped that findings from this study could serve to guide the development of a valid accent assessment technique that would offer clinical utility in the measurement of accents among international students attending U.S. colleges and universities. Reliable and valid measurement techniques could then be used to plan interventions- for accent reduction and document efficacy of treatment techniques.

## Method

### Participants

A convenience sample of 26 adult male speakers was selected to provide recorded speech samples for this investigation. Twenty-five of the speakers were from India and spoke American English as a second language; and one speaker was a college student from the mid-south region of the United States with no foreign accent. All of the international speakers spoke with an accent but had no personal experience with any formal therapeutic accent reduction programs.

Twenty-five normal hearing North American-born students for which English was their first language were selected to participate as listeners in this study. Listeners had no experience with accents or accent reduction and all reported no known hearing impairment. Twenty-two of the listeners were female and three were male. Of the original 25 listeners, two (8%) were dismissed based on their inability to make appropriate judgments of the extent of the accents and were replaced with listeners that could perform the tasks.

### Materials and Procedure

A 72 word passage with a complexity equivalent to an eighth grade reading level was used as the stimulus material to be recorded for this investigation. All speakers were allowed to familiarize themselves with the passage and to practice reading it aloud. Speakers were encouraged to relax and read the passage using

their natural rate, rhythm, and inflection patterns. Each speaker's reading of the passage was recorded individually and privately to eliminate any potential effects on reading skill related to reading with an audience. Digital recordings were made of each speaker's oral reading of the passage. An additional recording was made for one international speaker after he was asked to exaggerate his accent while reading the passage in his native language. Twenty-seven recorded passages were then transferred to a compact disc (CD).

Of the 27 recorded passages, 25 were test passages and 2 were considered validity passages. The two validity passages were used to ensure that listeners' responses were accurately reflecting their perception of the speaker's accent. Specifically, it was anticipated that for responses to be valid, listeners would provide responses indicating minimal or no accent for the one speaker with no foreign accent and responses indicating an extremely noticeable accent from the one international speaker exaggerating his accent while reading the passage in his native language.

### Procedure

An equal appearing interval (EAI) scale was used to obtain listener ratings of the perceived accent from each of the 27 recorded passages (25 test passages and 2 validity passages). The EAI scale ranged from 0 (no perceived accent) to 10 (extremely noticeable accent). Listeners were encouraged to rate the accents accordingly. If a listener felt that the speaker's accent was extremely noticeable they were instructed to rate the accent as a "10." At the opposite end of the scale, if the listener perceived the speaker to have no accent they were instructed to rate the accent as "0." The EAI 10 point scale allowed for subtle variations in scoring commensurate with the listeners' perceptions of the degree of accent. Listeners were encouraged to use any number on the EAI scale to within a decimal of .5 which they felt adequately reflected the extent of the speaker's accent.

In addition to the EAI scaling scores, direct magnitude estimations (DME) of the extent of the accents were obtained for each of the same recorded passages from the same listeners. During this phase, listeners were instructed to estimate the extent of the accent relative to a standard passage used as an anchor. The standard passage anchor was selected from the original pool of 25 test passages. This was accomplished by ranking the results from the EAI scaling phase on the basis of their median ratings and then selecting the passage at the 50th percentile as the standard passage

anchor. Listeners were informed that the standard passage anchor had been assigned a value of 100 and their task as listeners was to estimate the extent of the perceived accent from each of the test passages relative to the extent of accent of the standard passage anchor. To assist the listeners with their direct magnitude estimations a visual aid was employed with a standard-line anchor with an arbitrary value of 100 used to compare three separate lines of varying lengths. One line was half as long as the standard-line anchor and listeners were informed that a direct magnitude estimation of 50 would be an appropriate perceptual comparison. The second line was twice as long as the standard-line anchor and it was suggested that an estimation of 200 would be appropriate. Finally, a line with the exact same length as the standard-line anchor was depicted with the suggestion that an estimation of 100 would be appropriate. For each comparison, the standard passage anchor was played prior to the presentation of the 25 test passages and 2 validity passages. Additional validity of the estimations was obtained by having the listeners estimate the extent of the accent from the standard passage anchor when compared to itself.

Listeners were told that the investigation was designed to study accent. At no time was the word "intelligibility" used with listeners and care was taken, both in written and oral instructions, to avoid using terms indicating "more" and "less" except for the word "magnitude" during the magnitude estimation phase. All listening was performed in a modern amphitheater-type classroom equipped with a computer and high quality public address system capable of playing the recorded passages in any order. Care was taken to assure that the volume of passage was adequate and appropriate for comfortable listening. Listening practice was performed until the investigators were certain that listeners were familiar with the nature of the responses needed.

### Data Analysis

In order to address the validity of the DME results, a one-sample t-test was performed between the standard passage anchor (selected at the 50th percentile from the scaling phase and assigned a magnitude of 100) and the value of 100. Average values were computed from the listener's scaled scores for each of the 25 test passages for the EAI scaling procedure. Likewise, average values were computed for the DME phase for each of the speakers. A scatterplot and Pearson Product Moment coefficient of correlation was obtained between the average rating from each speaker and the average direct magnitude estimations from each speaker.



## Results

As noted in the previous section, 2 (8%) of the original 25 listeners that completed the EAI scaling and the DME portion of this investigation were dismissed from the study because of their inability to make appropriate judgments on the extent of the accents. One listener provided a scale judgment for one of the validity test passages indicating his/her scaled scores were not an accurate choice for what he/she perceived. Likewise, one listener provided a DME for one speech sample suggesting his/her inability to make the estimations. Both of these listeners were dismissed from the study and replaced with listeners who demonstrated adequate ability to complete the listening tasks.

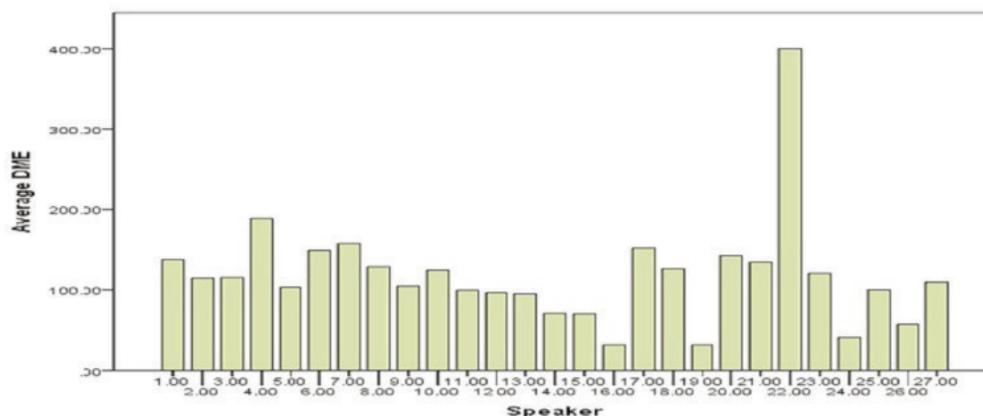
In order to verify that listeners were making valid EAI rating judgments, the average median rating scores were calculated from the 25 listeners for each of the 27 passages. Figure 1 is a bar graph depicting those results. Speaker 16 had no international accent. These results indicate that the listeners collectively rated his spoken passage as if he had little, if any, discernable accent. Similarly, speaker 22 was the international speaker who exaggerated his accent while reading the passage in his native language. It is clear from Figure 1 that the listeners judged his accent to be extremely noticeable and therefore gave his recorded passage an average median rating of 10. These results indicate that the listeners made EAI scale ratings that reflected the extent of the accent.

Average DME scores from the 25 listeners for each of the 27 passages were calculated and Figure 2 is a bar graph depicting those results. Again, speaker 16 was the speaker with no international accent and speaker 22 was the international speaker who exaggerated his accent while reading the passage in his native language. As with results from the EAI scaling method, it is clear from

Figure 1 that the listeners made valid DMEs reflecting the extent of the perceived accent.

For all subsequent statistical analyses, data for speaker 16 and 22 were removed. To further explore the validity of the DMEs for the standard passage, a one-sample t-test was performed on the average listener DME results for the standard passage and score of 100 (e.g. the standard passage anchor). The DME average listener score for the standard anchor passage was 96.047. Results of the t-test indicated no significant difference between the mean score for the standard passage anchor and 100 ( $t = -1.595, df=23, p>.124$ ) indicating that the listeners judged the standard passage anchor to be approximately 100. This finding suggests that it was appropriate for comparison purposes to select the speaker whose accent was judged to be the median as the anchor passage. Next a scatter plot was used to explore the relationship between the scores from the EAI method and the DME method. Results for the EAI scaling method were ranked from lowest to highest for the 25 speakers' passages. The DME results for the same speakers were plotted as a function of the EAI scaling procedure. An assessment indicates that average median EAI scale results increased at a similar rate as DME average scores. To further explore this relationship a Pearson Product Moment correlation coefficient was calculated between the pairs of average DME and EAI scores. The correlation was highly significant ( $r=.905, p>.000$ ) indicating that the relationship was positive and that a linear relationship existed between the two sets of scores. Figure 3 also shows the square of the correlation with a line of best fit added. Clearly, as the EAI scale scores increased there was a corresponding increase in the DME scores. The relationship between the DME and the EAI scaling was, therefore, linear.

Figure 1. A bar graph showing the average DME values for the 27 recorded passages.



### Discussion

The results from this study clearly indicated a linear relationship between DME and EAI ratings for extent of perceived accents. Applying the logic described by Stevens (1974) leads to the conclusion that either EAI or DME techniques would yield valid determinations of the extent of perceived accent and thus could be used for assessing the perception of extent of accent. The results indicate that the metathetic nature of the perception of accents was somewhat unexpected given the number of perceptual dimensions relative to other aspects of speech that had been shown to be prothetic (Berry & Silverman, 1972; Schiavetti et al., 1981; Toner & Emanuel, 1989).

As noted earlier, from a practical and clinical standpoint, ratings from an EAI scaling technique are easier to accomplish for at least two reasons. First, EAI scaling does not require a standard anchor or reference from which to base responses and second, EAI rating is a technique with which most listeners are familiar and can readily relate. Although not conclusive or alarming it should be noted that 1 (4%) of the original 25 listeners was unable to perform the EAI scaling technique despite explicit instructions and considerable practice. A possible explanation for this finding would be a supposition questioning the listener's motivation and commitment to the technique. Although there was ample evidence from this study to suggest EAI ratings represent a valid way to assess the perception of accents, there remains much to be learned about the reliability of the method. The experimental design of this study was not selected to focus on reliability per se, but there was nothing in the present findings to suggest the reliability of EAI ratings was suspect. Future research will be needed to explore the reliability of EAI ratings as they are applied to clinical assessment of the perception of the extent of accents.

Equal appearing interval scaling represents a relatively simple and practical method for obtaining perceptual measurements and could be easily adapted for routine clinical use in the study of accents. The procedure and design used in this study, although sound, does need further research before it is melded into a useful clinical tool. Specifically, the manner in which speech samples are harvested needs further investigation and refinement. Speakers in this study provided speech samples by reading aloud a short, novel passage in an unfamiliar setting. Currently, no research is available to indicate whether this method is the most appropriate format for harvesting natural examples of accent in speech of international students who speak English

as a second language. Future studies may need to explore the effects of oral reading in the non-native language on the speaker's accent. As a result, a more natural and less anxiety producing means to harvest speech samples for the purpose of studying accent may need to be developed. Also, having a cadre of listeners making accent judgments for every speaker is not clinically practical and highlighted a limitation of this study. It is not known, from this study, whether having listeners hear the same passage repeatedly for as many as 25 times created familiarity with the passage material and thereby influenced the perception of the extent of the accent. Future research might explore the validity and reliability of accent judgments made by a single listener both with and without professional experience with accent reduction therapy. Once these issues have been thoroughly addressed, further research can focus on determining the perceptual level of accent which is no longer perceived as a barrier to effective communication, thus guiding the practice of reasonable accent modification.

### Summary

There is little doubt that accent reduction will be a service increasingly in demand by international students enrolled in American universities. As the number of international students continues to rise in American universities, there will be a need to complement their education with other support services. One such service, when indicated and when called for, may be accent reduction.

The foundation of accent reduction will be an assessment instrument that is valid, reliable, evidenced based, norm referenced, and standardized. It is clear from the results of this study that a valid option for documenting the extent of an accent is equal appearing interval scaling. Work is still needed to address the reliability of the technique as it is applied to accent. Additionally, future research should be directed at a variety of international accents as well as the standardization of a strong clinical protocol.

### References

- American Speech Language and Hearing Association (2009). What is an accent? Retrieved from [http://www.asha.org/public/speech/development/accent\\_mod.htm](http://www.asha.org/public/speech/development/accent_mod.htm)
- Association of International Educators (2009). *NAFSA's guide to international student recruitment* (2nd ed.), Washington, DC.: NAFSA: Association of International Educators.



- Berry, R.C., & Silverman, F. H. (1972). Equality of intervals on the lewis-sherman scale of stuttering severity. *Journal of Speech and Hearing Research, 15*, 185-188.
- Fischer, K. (2009, November 16). Number of foreign students in U.S. hit a new high last year. Retrieved from <http://chronicle.com/article>
- Institute of International Education (2011). Publications and reports. Retrieved from <http://www.iie.org/en/Research-and-Publications/Publications-and-Reports>
- Marklein, M. B. (2011, November 14). More foreign students studying in USA. *USA Today*. Retrieved from <http://www.usatoday.com/news/education>
- Schiavetti, N., Metz, D. E., & Sitler, R.W. (1981). Construct validity of direct magnitude estimation and interval scaling: Evidence from a study of the hearing impaired. *Journal of Speech and Hearing Research, 24*, 441-445.
- Schiavetti, N., Sacco, P. R., Metz, D. E., & Sitler, R.W. (1983). Direct magnitude estimation and interval scaling of stuttering severity. *Journal of Speech and Hearing Research, 26*, 568-573.
- Shah, A. (2007). A comprehensive assessment of foreign-accented assessment speech: Test battery and model for assessment, course #3825. Retrieved from <http://speechpathology.com>
- Stevens, S. S. (1974). Perceptual magnitude and its measurement. In E.C. Carterett & M.P. Friedman (Eds.), *Handbook of perception* (Vol. 2). New York: Academic Press.
- Toner, M. A., & Emanuel, F.W. (1989). Direct magnitude estimation and equal appearing interval scaling of vowel and roughness. *Journal of Speech and Hearing Research, 32*, 78-82.
- Weismer, G., & Laures, J. S. (2002). Direct magnitude estimates of speech intelligibility in dysarthria: Effects of chosen standard. *Journal of Speech, Language, and Hearing Research, 45*, 421-433.
- Zraick, R. I., & Liss, J. M. (2000). A comparison of equal-appearing interval scaling and direct magnitude estimation of nasal voice quality. *Journal of Speech, Language, and Hearing Research, 43*, 979-988.

### **Acknowledgments**

This research was supported, in part, by a grant from the Arkansas State University Communication Disorders Foundation Account. The authors would like to thank Brandy Messer and Mary Hannah Pulliam for their assistance in preparing recordings and data collection.

---

### **About the Authors:**

*Dr. Mike McDaniel is a Professor of Communication Disorders in the Department of Communication Disorders at Arkansas State University. He has directed 15 master's thesis over the past 23 years. His research interests include auditory processing and speech intelligibility. He can be reached at [dmc-dan@astate.edu](mailto:dmc-dan@astate.edu).*

*Dr. Richard A. Neeley is a Professor and Chair, Department of Communication Disorders at Arkansas State University. His e-mail is [rneeley@astate.edu](mailto:rneeley@astate.edu).*

*Julie J. Isaacson is an Associate Professor of Nursing at the School of Nursing and Health Professionals, Arkansas State University. She can be reached at [jisaacson@astate.edu](mailto:jisaacson@astate.edu).*

*Dr. G. Daniel Howard is Interim Chancellor of Arkansas State University. His academic career spans more than 35 years. Among other things, Dr. Howard is an expert in globalizing institutions of higher education. He holds two master's degrees "with honors" and two doctoral degrees from Indiana University. He may be reached at [dhoward@astate.edu](mailto:dhoward@astate.edu).*