Intelligibility of British- and American-Accented Sentences for American Younger and Older Listeners With and Without Hearing Loss

Sadie Schilaty, AuD[1]   Sarah Hargus Ferguson, PhD[1]   Shae D. Morgan, AuD, PhD[1,2]
Caroline Champougny, MS[1]

[1] Department of Communication Sciences and Disorders, The University of Utah, Salt Lake City, Utah
[2] Department of Otolaryngology Head and Neck Surgery and Communicative Disorders, University of Louisville, Louisville, Kentucky

Address for correspondence Sarah Hargus Ferguson, AuD, sarah.ferguson@hsc.utah.edu

Abstract

Background   Older adults with hearing loss often report difficulty understanding British-accented speech, such as in television or movies, after having understood such speech in the past. A few studies have examined the intelligibility of various United States regional and non-U.S. varieties of English for American listeners, but only for young adults with normal hearing.

Purpose   This preliminary study sought to determine whether British-accented sentences were less intelligible than American-accented sentences for American younger and older adults with normal hearing and for older adults with hearing loss.

Research Design   A mixed-effects design, with talker accent and listening condition as within-subjects factors and listener group as a between-subjects factor.

Study Sample   Three listener groups consisting of 16 young adults with normal hearing, 15 older adults with essentially normal hearing, and 22 older adults with sloping sensorineural hearing loss.

Data Collection and Analysis   Sentences produced by one General American English speaker and one British English speaker were presented to listeners at 70 dB sound pressure level in quiet and in babble. Signal-to-noise ratios for the latter varied among the listener groups. Responses were typed into a textbox and saved on each trial. Effects of accent, listening condition, and listener group were assessed using linear mixed-effects models.

Keywords
► aging
► hearing loss
► speech perception
► accent

Results   American- and British-accented sentences were equally intelligible in quiet, but intelligibility in noise was lower for British-accented sentences than American-accented sentences. These intelligibility differences were similar for all three groups.

Conclusion   British-accented sentences were less intelligible than those produced by an American talker, but only in noise.

Older adults with hearing loss often report difficulty understanding speech in noise, and numerous studies confirm that noise has a disproportionately greater effect on speech understanding for older than younger adults.[1] Anecdotally, older adults with hearing loss also report difficulty understanding accented speech, including varieties of British English (BE) heard in television and movies. Considering Americans 50 years and older watch over 47 hours of television per week,[2] and one of hearing aid users' most frequent listening activities is listening to media,[3] understanding accented speech in media programs merits investigation. Beyond anecdotal complaints about understanding "Downton Abbey," difficulty comprehending accented speech has important implications in everyday life, especially in the context of globalization. Increased migration and the development of English as an international language[4] increase the likelihood of interactions wherein individuals speak different varieties of English. Previous research[5,6] has shown that older adults with and without hearing loss have difficulty understanding foreign-accented speech, especially in noise. However, no study to our knowledge has assessed the effects of different regional accents on speech understanding by older adults.

Different accents of a given language are characterized by variations in speech pronunciation. Perceptually, accents can be visualized along a scale with a listener's home accent on one end and foreign accents on the other.[7] Accents perceptually closer to a listener's home accent (e.g., regional dialects of a listener's native language) should be more intelligible and easier to process, while accents markedly further on the scale should be harder to understand. Support for this assumption is found in Bent et al,[8] where American listeners showed poorer identification of anomalous phrases spoken by Irish English talkers than of phrases produced by Spanish-accented talkers. Studies examining the intelligibility of various regional accents have also observed significant interactions between accent and listening condition. Clopper and Bradlow[9] mixed sentences produced by speakers of four American regional accents with speech-shaped noise at signal-to-noise ratios (SNRs) of +2, −2, and −6 dB, and found that the harder the listening condition, the bigger the intelligibility differences among the accents. Jacewicz and Fox[10] also found larger differences between two American regional accents at poorer SNRs than at better SNRs when listeners identified sentences in a background of two-talker babble. Finally, McCloy et al[11] demonstrated larger differences between two American regional accents when sentences were presented in speech-shaped noise at an SNR of +2 dB than at an SNR of +6 dB or in quiet, though the interaction was not tested statistically and may reflect ceiling effects in the easier conditions.

Several decades of research have shown that speech understanding differences between younger and older adults also depend on listening condition. Larger age effects occur in challenging listening conditions than in quiet, and the negative effects of distortions like noise,[1] reverberation,[12] and time compression[13] are greater in older adults than in younger adults, even after controlling for hearing sensitivity differences. In contrast, Ferguson et al[5] found the negative effect of foreign accent on word identification was approximately equivalent for young adults with normal hearing, older adults with essentially normal hearing, and older adults with hearing loss. Similar patterns seem to have occurred in other studies, although the effects of accent for individual groups are not always explicitly reported, and often interact with other factors, such as the presence or absence of sentence context[14] or different types of background noise.[15] While methodological differences make it difficult to compare results across studies, it appears that speech distortions have disproportionately greater negative effects in older adults only when those distortions have been applied to an existing speech signal, and not when the distortions occur during speech production. This potential contrast between externally applied and signal-inherent speech distortions could have implications for models of speech perception like Ease of Language Understanding (ELU),[16] which addresses the interaction between speech perception and cognition.

Despite differences among studies with regard to the effects of accent on older versus younger adults, it is evident that older adults with hearing loss show very poor speech understanding performance when the talker has a foreign accent, consistent with audiology patient reports. We designed the present preliminary study to assess whether American audiology patients' complaints of difficulty understanding British-accented English can also be demonstrated in the laboratory. Our hypotheses were (1) that listeners' word recognition accuracy would be poorer in British-accented sentences than in American-accented sentences, and (2) that as with foreign-accented speech, the size of any British accent effect would be the same for older and younger adults. To begin to examine whether any differences between younger and older adults were due to aging or to hearing loss, participants included older adults with essentially normal hearing as well as those with hearing loss.

Methods

Materials

The test stimuli for the experiment were two versions of the Basic English Lexicon (BEL) sentences.[17] The original BEL sentences were produced by one male and two female talkers speaking American English (AE). We also obtained a second version of the BEL sentences produced by a female talker speaking BE used in Pinet et al.[18] We chose the female AE talker whose speaking rate (4.28 syllables/second) more closely matched that of the BE talker (4.33 syllables/second). Each talker originally recorded 500 sentences, several of which were modified in some way for the BE recordings. Sentences containing different lexical items for the two talkers (e.g., film/movie, football/soccer) were excluded from the present experiment, yielding 364 sentences per talker. We then examined the remaining content of the original BEL sentence lists and chose four lists of 17 sentences (identical for the two talkers) for use in the experiment, as well as 10 additional unique sentences per talker for familiarization. Each sentence was padded with 50 ms of silence at

each end using Cool Edit 2000. Stimuli were then scaled to the same root mean square amplitude and resampled at 24,414 Hz for presentation via Tucker-Davis Technologies (TDT) audio hardware. A 12-talker AE babble (sampled from the Speech Perception in Noise test)[19] was used for the noise conditions.

Listeners

Three groups of listeners participated, all of whom were native AE speakers and denied any history of speech or language disorders. Fifteen young adults (ages 18–29, $\bar{x} = 22.7$ years), with normal hearing (YNH listeners) were recruited from the University of Utah Department of Psychology participant pool. Prior to the experiment, YNH listeners were determined by hearing screening to have pure-tone thresholds ≤ 25 dB hearing level (HL) from 250 to 8,000 Hz; they received course credit for participation. Twenty older adults with essentially normal hearing (ONH listeners; ages 67–81, $\bar{x} = 72.3$ years) and 19 older adults with hearing impairment (OHI listeners; ages 68–82, $\bar{x} = 73.1$ years) were recruited from a participant pool maintained by the second author. As members of this pool, ONH and OHI listeners had received a full audiological evaluation within 3 years of the present study. For the ONH group, we sought individuals with normal hearing (≤ 25 dB HL) for 250 to 4,000 Hz and no more than moderate hearing loss at higher frequencies, as in Ferguson et al.[5] To ensure a sufficiently large ONH group, however, we accepted six listeners with pure-tone thresholds of 30 dB HL at 4,000 Hz. OHI listeners were required to have mild to moderately severe sloping sensorineural hearing losses. Average audiograms appear in ►**Fig. 1**; the dashed line indicates the screening level for the YNH listeners. Most of the older adult listeners received payment for participating, though several chose to volunteer.
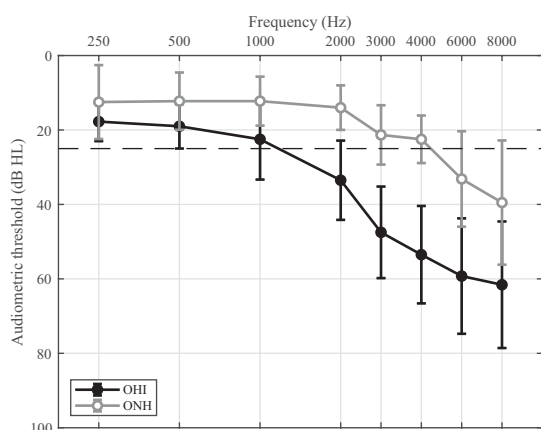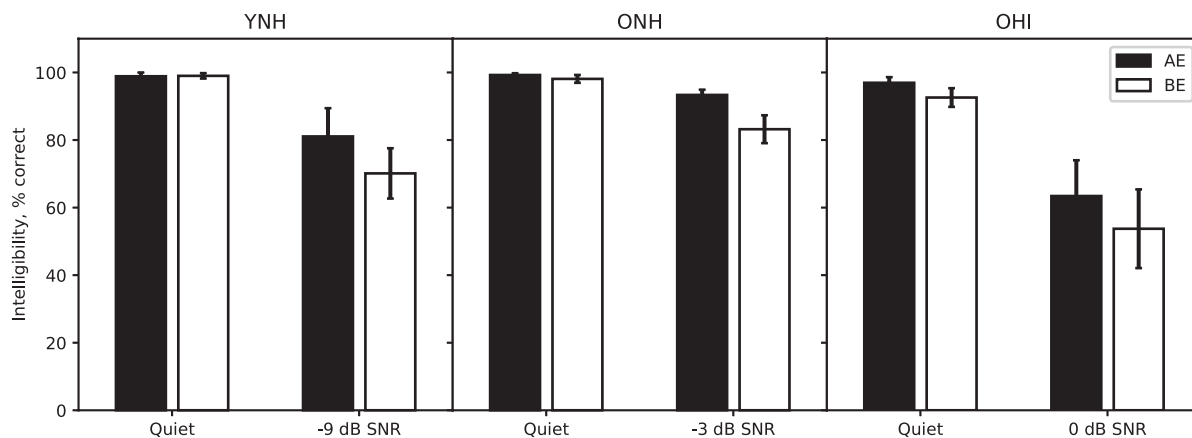
Procedures

All participants were tested individually in a single session lasting approximately 1 hour, seated facing a computer monitor, keyboard, and mouse in a quiet room. Listeners heard six blocks of sentences: 10 familiarization sentences (5 AE, 5 BE) in quiet, one list of AE in quiet, one list of BE in quiet, 10 familiarization sentences in noise, one list of AE in noise, and one list of BE in noise. Because our planned number of listeners per group would not permit the order of the four talker/listening condition combinations to be fully randomized, it was fixed for all listeners. The assignment of list to condition was pseudorandomized. After each block, the listener was encouraged to take a break, and the experimenter informed them about the content (i.e., which talker accent and listening condition to expect) of the upcoming block.

On each trial, a test sentence (in the quiet condition) or a test sentence and a segment of 12-talker babble (in the noise condition) were played from separate channels of a TDT RP2 real-time processor, attenuated by TDT PA-5 programmable attenuators to the desired overall level and SNR, mixed with a TDT SM5, and routed via a TDT HB7 headphone buffer to an insert earphone (E-A-RTONE 3A) for monaural presentation the ear that best matched the audiometric criteria for their listener group. If both ears matched those criteria, presentation was to the listener's preferred ear. The sentence presentation level was 70 dB sound pressure level for all listeners. Because we were particularly interested in the effects of accent, we wanted SNRs that yielded roughly equal performance on AE sentence recognition for all three groups. We also wished to avoid ceiling effects for AE sentences, and sought SNRs yielding AE sentence recognition scores of approximately 80% correct. Pilot testing identified these SNRs: –9 dB for YNH listeners, –3 dB for ONH listeners, and 0 dB for OHI listeners. Upon hearing each sentence, listeners either typed what they heard into a textbox or repeated what they heard to the experimenter, who typed their response. Listeners then pressed "enter" to trigger the next stimulus until the end of each sentence list. Listeners' responses were automatically recorded in a text file. Breaks were offered between test blocks.

Data Analyses

Each BEL sentence contains four keywords; thus, each list of 17 sentences contained 68 keywords. Participants' responses were scored by two experimenters. As in previously published studies,[10] keywords with spelling errors and homophones (e.g., *knight* vs. *night*) were counted as correct (worth 1), whereas those with additional or omitted phonemes were counted as wrong (worth 0). Percent correct scores were calculated for each listener and list and converted to rationalized arcsine units (RAUs).[20] RAU scores were analyzed using linear mixed-effects models performed in Stata 14 to test the effects of and interactions between three fixed factors: talker accent (AE, BE), listener group (YNH, ONH, OHI), and listening condition (quiet, noise). Listener was included as a random factor in all models.

Fig. 1 Mean audiometric thresholds for older adult listeners with hearing impairment (OHI; filled circles) and older adult listeners with normal hearing (ONH; open circles). Error bars indicate 95% confidence intervals. The dashed line corresponds to the screening level for the young adult listeners with normal hearing.

Fig. 2 Average percent correct sentence keyword recognition scores for three listener groups (YNH = young adults with normal hearing; ONH = older adults with normal hearing; OHI = older adults with hearing impairment), two talkers (AE = American English, BE = British English) in two conditions (quiet and noise; signal-to-noise ratio used for each listener group is shown in the *x*-axis label). Error bars indicate 95% confidence intervals.

Results

Average percent correct keywords for each listener group in each condition are displayed in ►**Fig. 2**, and ►**Table 1** summarizes the results of the statistical models. The linear mixed-effect models analyses revealed that the three fixed effects, talker, listening condition, and listener group, were all significant. Intelligibility was significantly higher for the AE talker than for the BE talker and significantly higher in quiet than in noise. Furthermore, overall intelligibility was lower for the OHI listeners than for the YNH or ONH listeners but did not differ between the YNH and ONH groups. Both of the two-way interactions involving listening condition were significant; listening condition interacted significantly with listener group and with talker. The two-way interaction

between talker and listener group, however, was not significant, nor was the three-way interaction.

Stratified analyses were then conducted to explore the bases of the significant two-way interactions; the results of these analyses are summarized in ►**Table 2**. Analyses for the condition × group interaction showed that the effect of listening condition was significant for all groups, but bigger for the YNH and OHI listeners than for the ONH listeners. The relationship among the three listener groups also differed for the two listening conditions. In quiet, the same pattern observed in the overall analysis held, with no difference between YNH and ONH listener performance ($\bar{x} = 117.7$ and 115.4 RAU), but significantly lower OHI listener performance ($\bar{x} = 104.9$ RAU)

Table 1 Summary of statistics for the linear mixed-effects models

Main effects	β	z	p
Talker	9.07	5.10	< 0.001
Condition	36.01	20.25	< 0.001
OHI vs. YNH	18.35	5.47	< 0.001
OHI vs. ONH	23.67	6.58	< 0.001
YNH vs. ONH	5.31	1.42	0.156
Interaction effects			
Listener group x Condition		4.45	< 0.001
Talker × Condition		2.54	< 0.02
Listener group × Talker		0.88	0.337
Listener group × Talker × Condition		< 1.3	> 0.2

Abbreviations: OHI, older adults with hearing impairment; ONH, older adults with normal hearing; YNH, young adults with normal hearing.
Note: Main effects and interaction effects for listener group (YNH; ONH; OHI), condition (in noise vs. in quiet), and talker (AE, American English; BE, British English) are presented.

Table 2 Summary of statistics for stratified analyses of interaction effects

Condition × Group	β	z	p
Condition effect, YNH	−40.43	−12.88	< 0.001
Condition effect, ONH	−23.10	−10.66	< 0.001
Condition effect, OHI	−45.88	−13.15	< 0.001
YNH vs. ONH, Quiet	−2.30	−0.94	0.349
OHI vs. YNH, Quiet	12.90	2.45	< 0.001
OHI vs. ONH, Quiet	10.60	2.34	< 0.001
YNH vs. ONH vs. OHI, Noise	> 13.00	> 2.25	< 0.05
Talker × Condition			
Talker effect, Noise	−13.05	−5.1	< 0.001
Talker effect, Quiet	−5.09	−2.85	0.004
Condition effect, BE	−39.99	−14.05	< 0.001
Condition effect, AE	−32.03	−10.47	< 0.001

Abbreviations: OHI, older adults with hearing impairment; ONH, older adults with normal hearing; YNH, young adults with normal hearing.
Note: Interactions between listener group (YNH; ONH; OHI), condition (in noise vs. in quiet), and talker (AE, American English; BE, British English) are presented.

than in both groups with normal hearing. In noise, significant differences were found between all three groups, with ONH listeners showing the highest performance ($\bar{x} = 92.3$ RAU), followed by YNH listeners ($\bar{x} = 77.2$ RAU), and lowest performance in OHI listeners ($\bar{x} = 59.1$ RAU).

Stratified analyses of the talker × condition interaction showed that the effect of talker was significant in both listening conditions, but the difference between the talkers was larger in noise than in quiet. The negative effect of noise was sizeable and significant for both talkers, but slightly larger for the BE talker than the AE talker.

Discussion

Consistent with our first hypothesis, keyword recognition accuracy was significantly lower for British-accented sentences than for sentences produced by an American talker. However, the magnitude of the talker effect was quite small, especially in quiet. Indeed, the difference between the talkers in ►Fig. 2 is negligible in quiet, with performance close to 100% correct for the YNH and ONH listeners and well above 90% correct for the OHI listeners for both talkers. The magnitude of the accent effect in the present study was bigger when sentences were presented in 12-talker babble, but still only approximately 10 percentage points for all three listener groups. This is comparable to the 15-percentage-point difference Jacewicz and Fox[10] found between two accents for sentences presented to YNH listeners in two-talker babble at an SNR of –3 dB, but smaller than the range observed across four American regional accents when YNH listeners identified sentences in speech-shaped noise at an SNR of +2 dB (∼25 percentage points) in Clopper and Bradlow.[9] The effect of foreign accent appears to be slightly larger than that of regional accents, with a word intelligibility difference between native English and native Spanish speaker(s) across conditions and listener groups of 30 to 35 percentage points in Ferguson et al[5] and in Gordon-Salant et al.[6]

The present data are also consistent with our second hypothesis, that the effect of British accent would be the same for the younger and older adults. Just as in the Ferguson et al[5] study of foreign accent, there was no significant interaction between talker accent and listener group. These results stand in sharp contrast with an extensive literature showing that noise and other distortions have a disproportionately greater negative effect on speech perception performance for older adults than younger adults, even when signal audibility is equated.[1,12,20] It thus appears that the aging brain copes with signal-inherent distortions much like the younger brain does, but has a much harder time coping with externally applied distortions. While this contrast would have implications for models of speech perception and cognition like ELU,[16] it is important to note that only a few studies have specifically evaluated the interaction between talker accent and listener age. More research is needed before a definitive statement can be made about how older adults are affected by different speech signal distortion types.

It will be important that this future research avoid the methodological limitations of the present preliminary study.

For example, our test materials included just one talker per accent. While studies of foreign-accented speech perception have used only one or at most two native or non-native talkers,[5,15] most studies of regional accent effects for YNH listeners have used larger numbers of talkers for each accent. For example, Jacewicz and Fox[10] had four talkers each from two accent regions, while Clopper and Bradlow[9] had six talkers each from four regions. Future studies of the perception of accented speech by older adults should certainly include more talkers, although the ideal number is unknown, and the number of talkers must be weighed against other methodological concerns, including materials, listener groups, and statistical approach. Future studies could also examine the speech materials from different accents in greater depth to determine how similar or dissimilar each accent is from the others. This could be done perceptually, using subjective ratings; acoustically, measuring both segmental and suprasegmental characteristics; or ideally via some combination of both methods.

The use of different test SNRs for the three groups in the present study was problematic. While pilot tests showed these SNRs produced equal performance among the three groups for the AE talker, actual results reported here show that our pilot sample did not accurately represent the performance of the larger sample. Unfortunately, the equal performance at differing SNRs did not occur in the actual experiment, suggesting that we assessed the groups at different points in the psychometric function, which should be noted for the careful interpretation of our results. In addition, it is possible the fixed order of the four talker/listening condition combinations, where the BE listener in noise was the last condition, exaggerated the differences between the conditions and possibly the talkers due to listener fatigue effects. Future experiments could address these problems by testing all listeners at multiple SNRs and fully randomizing the test conditions.

Conclusion

In summary, the present study supports previous work suggesting that stimulus-inherent characteristics like talker accent seem to have similar effects on younger and older adults, in contrast with externally applied distortions like noise which have a disproportionately negative effect on older adults with hearing loss. Specific to this work, the BE talker was less intelligible to American listeners when presented in noise, but equally intelligible in quiet. Accent effects were consistent across age (i.e., for ONH and YNH listeners) or hearing status (i.e., for ONH and OHI listeners). The reasons underlying the difference between "internal" (stimulus-inherent) versus "external" distortions are an interesting area for future research.

References

1 Dubno JR, Dirks DD, Morgan DE. Effects of age and mild hearing loss on speech recognition in noise. J Acoust Soc Am 1984;76(01):87–96

2 Nielsen The total audience report: Q1. 2017. Available at: www.nielsen.com. Accessed September 14, 2020

3 Hasan SS, Chipara O, Wu Y-H, Aksan N. Evaluating auditory contexts and their impacts on hearing aid outcomes with mobile phones. Pervasive Health '14: Proceedings of the 8th International Conference on Pervasive Computing Technologies for Health Care. 2014:126–133

4 Sharifian F. English as an International Language: Perspectives and Pedagogical Issues. Bristol: Multilingual Matters; 2009

5 Ferguson SH, Jongman A, Sereno JA, Keum KA. Intelligibility of foreign-accented speech for older adults with and without hearing loss. J Am Acad Audiol 2010;21(03):153–162

6 Gordon-Salant S, Yeni-Komshian GH, Fitzgibbons PJ. Recognition of accented English in quiet and noise by younger and older listeners. J Acoust Soc Am 2010;128(05):3152–3160

7 Bent T, Atagi E, Akbik A, Bonifield E. Classification of regional dialects, international dialects, and nonnative accents. J Phonetics 2016a58:104–117

8 Bent T, Baese-Berk M, Borrie SA, McKee M. Individual differences in the perception of regional, nonnative, and disordered speech varieties. J Acoust Soc Am 2016b140(05):3775–3786

9 Clopper CG, Bradlow AR. Perception of dialect variation in noise: intelligibility and classification. Lang Speech 2008;51(Pt 3):175–198

10 Jacewicz E, Fox RA. The effects of dialect variation on speech intelligibility in a multitalker background. Appl Psycholinguist 2015;36:729–746

11 McCloy DR, Wright RA, Souza PE. Talker versus dialect effects on speech intelligibility: a symmetrical study. Lang Speech 2015;58 (Pt 3):371–386

12 Gordon-Salant S, Fitzgibbons PJ. Comparing recognition of distorted speech using an equivalent signal-to-noise ratio index. J Speech Hear Res 1995;38(03):706–713

13 Wingfield A, McCoy SL, Peelle JE, Tun PA, Cox LC. Effects of adult aging and hearing loss on comprehension of rapid speech varying in syntactic complexity. J Am Acad Audiol 2006;17 (07):487–497

14 Gordon-Salant S, Yeni-Komshian GH, Fitzgibbons PJ, Cohen JI. Effects of age and hearing loss on recognition of unaccented and accented multisyllabic words. J Acoust Soc Am 2015;137(02): 884–897

15 Gordon-Salant S, Yeni-Komshian GH, Fitzgibbons PJ, Cohen JI, Waldroup C. Recognition of accented and unaccented speech in different maskers by younger and older listeners. J Acoust Soc Am 2013;134(01):618–627

16 Rönnberg J, Rudner M, Foo C, Lunner T. Cognition counts: a working memory system for Ease of Language Understanding (ELU). Int J Audiol 2008;47(Suppl 2):S99–S105

17 Calandruccio L, Smiljanić R. New sentence recognition materials developed using a basic non-native English lexicon. J Speech Lang Hear Res 2012;55(05):1342–1355

18 Pinet M, Gan Y, Evans B, Iverson P. Intelligibility of British English accents in noise for second-language learners. Proceedings of ICPhS 2015

19 Kalikow DN, Stevens KN, Elliott LL. Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. J Acoust Soc Am 1977;61(05): 1337–1351

20 Studebaker GA. A "rationalized" arcsine transform. J Speech Hear Res 1985;28(03):455–462