

Effects of Stress, Stop Release, and Familiarization on Speech Recognition Thresholds

Noah Eggebraaten¹ and Youkyung Bae^{2*}

¹Mayo Clinic Health System, Albert Lea, Minnesota, USA

²Department of Speech and Hearing Science, The Ohio State University, USA

*Corresponding author: Youkyung Bae, The Ohio State University, USA, Tel: 614-688-4948; Fax: 614-292-7504; E-mail: bae.180@osu.edu

Received date: December 01, 2016; Accepted date: January 12, 2017; Published date: January 13, 2017

Copyright: © 2017 Eggebraaten N. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Objective: This study examined outcomes of common procedural variations of speech recognition threshold (SRT) testing, specifically related to the effects of equal syllable stress, word-final stop consonant release, and prior-familiarization, with the participants' language status taken into account.

Methods: SRTs were obtained from 40 adults with normal hearing. Twenty participants received prior-familiarization with the spondee list and the other 20 received no prior-familiarization. Repeated SRT tests were administered using three different recordings which varied in syllable stress and word-final stop release patterns.

Results: The group with prior-familiarization demonstrated a threshold that was significantly lower than the group without prior-familiarization, by approximately 5 dB HL. Despite the statistically significant effects of equal syllable stress and word-final stop release on SRTs, the magnitude of SRT changes elicited by these acoustic-phonetic variations was only slightly above 1 dB HL. The monolinguals generally outperformed the bilinguals in SRT outcomes with the threshold difference less than 3 dB HL.

Conclusion: Findings from the present study suggest that familiarizing listeners with test vocabulary prior to SRT administration should continue to remain an important procedural requirement. Future research addressing the extent to which acoustic-phonetic variations of spondee production affect SRTs in individuals with hearing impairments is warranted.

Keywords: SRT; Speech reception; Familiarization; Stress; Final stop release; Acoustic-phonetic variation; Bilingual speech reception

Introduction

In audiology, pure-tone audiometry is often considered as the primary tool of clinicians, but Martin and Clark [1] write that "the hearing impairment inferred from a pure-tone audiogram cannot depict beyond the grossest generalizations, the degree of disability in speech communication caused by hearing loss" (p. 126). Audiologists have in turn thought it fitting to use speech stimuli to test a patient's ability to understand the spoken word, which has placed the speech recognition threshold (SRT) among the standard battery of tests used to evaluate hearing. Given its strong correlation with the pure-tone average (PTA), the SRT is routinely used as a validation check for the PTA.

Establishing a standard measure for the SRT has been a long-standing interest to researchers and clinicians. The development of a standardized metric used by a majority of audiologists was previously pursued to improve test validity and reliability within and between clinicians and clinics [2]. However, given that variability in speech production includes, but is not limited to, phonetic makeup, prosodic tendencies of the speaker, and suprasegmental features, the difficulty of developing and implementing standard spoken test materials and protocols is considerable and continues to affect current practices. As a result of several attempts to create spoken test materials [3-5], Hudgins

and colleagues [4] provided testing criteria, most notably a standardized list of individual words (spondees). The essential characteristics for the development of a spondee word list were based on the desire to control for acoustic and psychometric variables. Researchers agreed that an ideal word list would include words that are familiar to the listener, phonetically dissimilar, homogeneous with respect to audibility, and that feature a normal sampling of English speech sounds [4]. While several researchers disagree how closely the current spondaic word lists actually adhere to the criteria for test material proposed by Hudgins and colleagues, these criteria continue to be represented in the Central Institute for the Deaf Auditory Test W-1 (C.I.D. W-1) word list, one of the most widely investigated spondee lists.

Hirsch et al. [5] conducted a series of experiments containing several word types, including spondees, in hopes of developing new tools of measurement in speech audiometry, as well as improving upon existing measures. The goal was to create a spondaic list of words with similar properties that adhered to the criteria for test material suggested by Hudgins and colleagues. The initial word list began with 84 spondees, which were rated for familiarity based on a three-point scale. Among those words rated as most commonly heard or familiar to listeners, outlying words, which were either too easy or too hard to hear, were further eliminated from the list in order to ensure equal intelligibility of the list. The resulting "homogeneous" list, known as the C.I.D. W-1, consisted of 36 spondees and became the standard word list used for adults when obtaining the SRT, and is now the

standard list prescribed by the American Speech-Language-Hearing Association. The ASHA guidelines for determining the SRT define it as “the minimum hearing level for speech at which an individual can recognize 50% of the speech material” [2].

Martin and Clark [1] described the recommended clinical preparation of ASHA’s prescribed method for determining the SRT with the following steps: 1) familiarize the listener with the spondaic words in the word list to be used; 2) ensure that the vocabulary is familiar; 3) establish that each word can be recognized auditorily; and 4) ascertain that the patient’s responses can be understood by the clinician (p. 132). Steps 1 and 2 ensure that the closed set of spondees is familiar to each patient and to ensure that no single word presents difficulty to the patient [1]. Steps 3 and 4 ensure that the audiologist can properly record the patient’s responses. Steps 3 and 4 are intuitively necessary, but Steps 1 and 2, related to familiarizing the patient with the spondee list, have not made their transference to the audiology clinic with great success. Several questionnaires conducted in North America have indicated that only about 43%-58% of audiologists regularly or always familiarize patients prior to SRT testing [6-8].

The additional time required to familiarize patients with the test material has been cited as rationale for skipping the familiarization process [9]. Findings, however, showed that the process of familiarization resulted in a 20% average increased number of correct responses, indicating that familiarization aids in the ease of recognition of most spondees. Wilson and Margolis [10] analyzed the data presented by Conn et al. [9] and specifically pointed out that participant responses were, as a whole, more homogeneous with respect to intelligibility when prior familiarization was used for the 36-word list. In other words, without being adequately controlled, the degree of the participant’s familiarity with the testing procedure and test vocabulary may confound with his/her true auditory threshold value. Empirical evidence was further suggested that prior knowledge of spondee test vocabulary could improve thresholds by 4-5 dB [11]. In that study, the group that received repeated testing with the same set of test vocabulary demonstrated the most noticeable performance improvement followed by the group that received repeated testing with different sets of test vocabulary. Not surprisingly, the group that had prior knowledge of test vocabulary (i.e., by repeating each word aloud) outperformed the other groups. It is noteworthy that the SRT outcomes from this group with prior knowledge remained fairly consistent in repeated testing conditions within a 0.3 dB between-sessions difference. Another study [12] reported similar findings, in which the group with prior knowledge (i.e., highly experienced and familiar with the testing procedure and test vocabulary) significantly outperformed the group without prior knowledge. The magnitude of SRT change across repeated tests or sessions, however, was on the order of only a few tenths of a dB. Taken together, both of the aforementioned studies [11,12] suggest that SRT outcomes, in response to the repeated testing condition, continue to improve as participants are being familiarized with testing procedures and test vocabulary. Meanwhile, gains through learning or familiarization appear to saturate; that is, once familiarized with the procedure and test vocabulary, participants gain little additional improvement in response to successive testing conditions.

Spondees, by definition, must be “two-syllable words with equal stress on both syllables” and are typically represented by compound words such as “baseball” [2]. To account for stress discrepancies (i.e., primary versus secondary stress) found in most disyllabic English words, audiologists routinely use calibrated volume unit (VU) meters

to ensure that both syllables peak at zero on the VU meter [1,13]. Stress, however, is associated not only with amplitude, but also syllable duration and fundamental frequency (f_0) [14,15], with the latter two cited as the dominant features [16].

Bettagere [13] analyzed stress patterns in a professional recording (Q/Mass Speech Audiometry, Qualitone, 1988) of a spondaic list intended for clinical use. Results from the acoustic analysis demonstrated statistically significant differences between syllables in the measures of duration and f_0 , but not in the measure of amplitude, presumably due to the use of the volume unit (VU) meter monitoring peak syllable levels. Further, results from the perceptual analysis showed that, upon listening to the recording in question, over one half of the participant responses indicated the perception of unequal stress. It was concluded that the professional recording in question did not demonstrate equal syllable stress acoustically or perceptually and that both f_0 and duration should be considered as controls for syllable stress when determining the SRT. Further research investigating possible effects of balanced and unbalanced syllable stress on SRT outcomes was suggested. Given that stress is associated with salience of a syllable and that the SRT measures the recognition of words, variation of stress pattern within and between clinicians during the production of spondees could conceivably affect the validity of an SRT test, as the salience of syllables might not be uniformly presented at all decibel levels. Providing that stressed syllables appear to aid in faster phoneme recognition than unstressed syllables [17], it can be hypothesized that greater stress given to either syllable in SRT testing would likely aid in identification of that syllable. This, in turn, could violate the ASHA protocol [2] requiring equal stress for each spondaic syllable.

Along with syllable stress, another variable not addressed in ASHA’s 1988 guidelines is the optional release of word-final stop consonants. The brief noise burst (aspiration) accompanying voiceless stop consonants is known to provide acoustic information for a listener to determine the place of articulation [18]. In fact, unreleased voiceless stops (/p/, /t/, /k/) are known to provide a considerably lower level of acoustic information necessary for place identification than their released counterparts [19]. Similar results were found for voiced stops (/b/, /d/, /g/), albeit to a lesser degree. Lisker [20] also demonstrated that unreleased voiceless stops in final position are generally less intelligible than their released counterparts. There are 17 spondees in the C.I.D. W-1 list that end in a stop consonant, which have the option to be released or unreleased when produced in normal conversational English. Not controlling for the release status of these consonants may provide patients with altering signals of the test material when obtaining the SRT. This could be especially true in the monitored live-voice method, which appears to be the primary method used by 94% of the licensed audiologists in the USA [8].

In addition to the aforementioned variables, a growing body of literature has suggested the effect of bilingualism on speech audiometry. Non-fluent English speakers typically demonstrate poorer performance on instruments using English stimuli. Ramkissoon et al. [21] obtained SRT scores using the C.I.D. W-1 list from a group of listeners with limited English proficiency and found that the thresholds did not correlate well with PTA scores and were significantly different from thresholds obtained from native English speakers. However, effects of bilingualism in speakers proficient in English show less straightforward effects. For instance, Mayo, Florentine, and Buus [22] found that sequential Spanish to English bilinguals who learned English before age six had better English speech perception in noise than bilinguals who learned English after age fourteen. Interestingly,

the difference between bilinguals and monolinguals seems to be diminished when these listeners perceive speech in quiet environments [22,23]. It is less well understood how presentation of English words at low decibel levels (e.g., SRT testing) affects bilinguals, particularly proficient English speakers, compared to monolinguals.

To date, acoustic-phonetic variations in spondee production, especially related to syllable stress and word-final stop consonant release patterns, have received little attention. Despite considerable variability in spondee production across clinicians and the potential theoretical impact on SRT outcomes [13,19,20], no empirical data have been reported in the literature. Meanwhile, prior familiarization, with its well-known effect on SRT outcomes [11,12], has been considered an essential step of the ASHA's guidelines [2]. Nonetheless, the reported rate of guideline adherence to prior familiarization remained below 60% among audiologists in North America [6-8]. Taken together, procedural variations clearly appear to exist in SRT administration, and yet, it is not clear to what extent these commonly observed procedural variations should be considered acceptable or unacceptable in practice. If acoustic-phonetic variations of spondee production, along with the patient's prior familiarization of test vocabulary, alter the SRT in a clinically meaningful way, it is conceivable that a "worst case scenario" could exist. In this scenario, all three (no prior familiarization, unequal syllable stress, and unreleased word-final stop consonants) could aggregate, hypothetically causing poorer thresholds. In an effort to further clarify these issues, the present study examined individual effects of prior familiarization, equal syllable stress, and word-final stop release on SRTs, while taking the listeners' language status into account.

Methods

Participants

A total of 40 healthy adults with normal hearing were recruited from a local university setting through convenience sampling. The recruitment site and the sampling method resulted in a disproportionate gender representation (3 males; 37 females) of the participants from a relatively homogeneous age range between the ages of 21 and 33 years (mean: 24 years). Any participant who did not demonstrate a normal PTA between -10 and 15 dB HL according to the ANSI-2004 scale was dismissed from the study. Only those who spoke American English as their first language or considered themselves as having native-like proficiency in English were included. Participants were recruited in the southern New Mexico region, where approximately 36% of the population speaks a language at home other than English [24]. Twenty participants reported being monolingual, 18 reported being bilingual, and two reported being trilingual. Spanish was the most commonly spoken language in addition to American English. This study was approved by the University Institutional Review Board, and informed consents were reviewed and signed by individual participants.

Test stimuli

Using the C.I.D. W-1 list of 36 spondee words (Table 1), three sets of test stimuli targeting the acoustic-phonetic variables for stress and word-final stop release were recorded by a male speaker (N.E.) using Pro Tools recording software (Avid Technology Inc., Burlington, MA). To best approximate a common clinical environment where monitored live-voice testing is used, all recordings were monitored using a VU meter as a visual control to ensure that the peak volume of each

syllable did not exceed 0 dB. Prior to recording, a 1000 Hz pure-tone was first calibrated to peak at 0 dB on the Pro Tools VU meter, then a microphone (Blue Yeti Pro, Blue Microphones, Westlake Village, CA) was calibrated so that peak syllable levels matched the calibrated pure-tone signal without surpassing the specified decibel level on the Pro Tools VU meter. The files were then imported to recording software (Garageband, Apple Inc., Cupertino, CA), where peak volume levels received visual waveform inspection and volume adjustment through the interface so that each syllable in each spondee peaked at 0 dB on the calibrated GSI 61 Clinical Audiometer (Grason-Staler, Milford, NH). The spondees were exported to iTunes (Apple Inc., Cupertino, CA) and played through the audiometer during the experiment.

Baseball	Airplane	Iceberg	Sunset
Hardware	Armchair	Playground	Stairway
Woodwork	Workshop	Birthday	Eardrum
Doormat	Northwest	Railroad	Grandson
Sidewalk	Farewell	Mousetrap	Mushroom
Whitewash	Horseshoe	Hotdog	Oatmeal
Pancake	Cowboy	Daybreak	Toothbrush
Drawbridge	Greyhound	Inkwell	Schoolboy
Padlock	Duckpond	Headlight	Hothouse

Table 1: The C.I.D. W-1 list of 36 spondee words.

First recording set (R1): Spondees unequally stressed with unreleased word-final stops

The first recording (R1) was designed to contain unequal syllable stress, as well as unreleased word-final stop consonants on all qualifying spondees. A possible propensity was noted in which clinicians tend to increase the f_0 and duration of the second syllable in an effort to produce spondees with equal stress, while maintaining relatively equal amplitude [13]. Thus, "unequal stress" acoustically assumed higher f_0 and longer duration on the second syllable of each spondee than on the first syllable in the creation of R1. Perceptually, "unequal stress" assumed that one investigator (N.E.) and a licensed speech-language pathologist (as an external rater) independently and audibly perceived unequal stress in each word on the R1 list. Inclusion of a perceptual component in the working definition of unequal stress was necessary, given that subtle acoustic inequalities may not be sufficient for a listener to perceive two syllables as being unequal [13].

The variable of "unreleased status" of word-final stops in R1 also required acoustic and perceptual guidelines. Acoustically, the unreleased status was determined by identifying no final release burst associated with released stops on spectrograms and waveforms of each word in the R1 list containing a word-final stop consonant. Perceptually, the unreleased status assumed that the investigator and the speech-language pathologist independently perceived no audible release of word-final stop consonants in R1.

Second recording set (R2): Spondees equally stressed with unreleased word-final stops

The second recording (R2) was designed to contain equal syllable stress in each spondee, as well as unreleased word-final stop

consonants on all qualifying spondees. Similar to R1, a working definition of “equal stress” assumed that the investigator and the speech-language pathologist independently and audibly perceived equal stress between each word on the R2 list. Acoustically, “equal stress” assumed that the f_0 and duration for both syllables in R2 were less than the lowest value of any second syllable in R1. This seemed appropriate given that the second syllable in R1 was relatively stressed compared to the first syllable in R1, and both syllables in R2 were designed to contain a lesser f_0 and duration than the intentionally stressed syllables in R1.

Third recording set (R3): Spondees equally stressed with released word-final stops

The third recording (R3) was designed to contain equal syllable stress in each spondee, as well as release of word-final stop consonants on all qualifying spondees (17 out of 38 spondees in the C.I.D. W-1 list). Those spondees containing word-final stops followed the same requirements of “equal stress” as outlined for R2 but with a caveat regarding their duration. As voice onset time following aspiration in adults can last as long as 90 milliseconds (e.g., /k/) [25], faster production of the spondee would be necessary to meet the requirement for equal stress mentioned above in terms of duration. Therefore, the researchers measured the duration of spondees containing released word-final stops up until the end point of the sound preceding the word-final stop consonant, as interpreted in each spectrogram and waveform. This measure maintained a relative uniformity in spondee production for the entire R3 list, while isolating the released status as a variable of interest. Acoustically, the researchers considered spondees containing word-final stops as released by identifying a final burst of noise in the spectrogram and waveform. Perceptually, the released status assumed that the investigator and the speech-language pathologist independently perceived audible release of final stop consonants in the R3 list.

Acoustic controls of test stimuli

While the VU meter served as a control for amplitude during the creation of the test stimuli, special attention was also required for the stress variables of duration and f_0 . The speaker utilized metronome

beats and piano notes that were simultaneously delivered through the headphone as references for the syllable duration and f_0 . The metronome beat was set at 200 beats per minute (bpm), meaning that one beat occurred every 300 msec. During the creation of R1, the speaker attempted to produce spondees with unequal syllable stress, consisting of two beats (600 msec) and three beats (900 msec) for the first and the second syllable, respectively. During the creation of R2 and R3, the speaker attempted to produce spondees with equal syllable stress consisting of two beats (600 msec) in length for each syllable. Piano notes, which approximated the speaker’s own habitual f_0 , were chosen to provide consistent and comfortable production of the test stimuli. In R1, the notes G2# (103.83 Hz) and A2# (116.54 Hz) respectively served as references for the first and second syllables. This helped ensure unequal f_0 s between syllables while maintaining perceptual distinctness and normal variability of pitch during conversation [26]. In R2 and R3, G2# served as a reference for the first and second syllables to ensure uniform f_0 s.

To ensure that each recording conformed to its assigned acoustic-phonetic characteristics in relation to the syllable stress and word-final stop release pattern, acoustic analyses were performed using PRAAT software [27]. Table 2 provides summary statistics of the syllable duration (msec), f_0 (Hz), and intensity (dB) measures for R1, R2, and R3. The syllable duration refers to the entire duration of each syllable with the exception of syllables beginning or ending with stop consonants. For a syllable beginning with a stop consonant, the stop gap segment was excluded from the duration measure. Additionally, for a syllable ending with a stop consonant, the entire stop was excluded from the duration measure. The intensity measure refers to the average intensity of the same speech segments used for measuring the syllable duration. The f_0 measure refers to the average f_0 measured based on the syllable nucleus. To estimate inter-rater agreement on acoustic measurement of spondees, about 30% of the spondee stimuli from each recording were re-measured. Agreement between two sets of acoustic measurements was assessed using two-way mixed, absolute agreement, and single measures A Intraclass Correlation Coefficient (ICC). The ICC coefficients of 0.996, 0.998, and 0.956 were observed for the duration, f_0 , and intensity measures, respectively, suggesting a high level of agreement between the two sets of acoustic measurements.

	R1		R2		R3							
	Unequal stress	syllable	Unreleased final stop	Equal stress	syllable	Unreleased final stop	Equal stress	syllable	Released final stop			
	1st Syllable		2nd Syllable		1st Syllable		2nd Syllable		1st Syllable		2nd Syllable	
Syllable duration (msec)	565.49 (79.83)		960.88 (81.11)		600.00 (67.89)		569.31 (92.20)		559.29 (72.49)		562.58 (79.82)	
f_0 (Hz)	565.49 (79.83)		116.57 (1.90)		104.78 (2.68)		104.23 (1.21)		104.31 (1.54)		103.92 (1.56)	
Intensity (dB)	59.60 (1.55)		61.74 (1.11)		59.83 (2.10)		59.97 (1.59)		59.16 (1.95)		59.02 (1.53)	

Table 2: Means and standard deviations (in parentheses) of the acoustic measures including syllable duration, fundamental frequency (f_0), and intensity of the first and second syllables across three different recording conditions (R1, R2, and R3).

Procedure

To ensure that each participant had normal hearing, pure-tone thresholds were obtained using a descending/ascending method described by Martin & Clark [1]. Pure-tone averages (PTAs) were

determined based on the thresholds at frequencies of 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz. As outlined in ASHA’s guideline [2], SRTs were obtained using the 2 dB descending technique with two sets of randomized 36 words of the C.I.D. W-1 list. The word lists were played subsequently on repeat during the SRT testing of each respective

recording. One spondee was presented at a time. Eighteen SRT outcomes were determined by administering each recording (R1, R2 and R3) until six SRT scores were established per recording (3 recording sets × 2 ears × 3 repetitions).

The SRT testing had two procedural versions (Table 3). Forty participants were randomly assigned to two different versions of testing. Twenty participants (12 monolinguals and 8 bilinguals) received Version A, and 20 participants (8 monolinguals and 12 bilinguals) received Version B. The participants assigned to Version A were familiarized with the list of spondees prior to administration of SRT testing by reading the list of 36 spondees aloud to the researcher. Any inaccurate readings of the spondees were corrected by the researcher and then repeated by the participant. The SRTs were then obtained for R1, R2, and R3, respectively. In Version B, SRT testing using R1 was administered prior to familiarization. Participants assigned to Version B were familiarized with the list of spondees after administration of SRT testing with R1, and then received testing for R2 and R3. These two versions provided data regarding the effects of being familiarized to a spondee list prior to testing versus receiving no prior familiarization. Although randomizing the order of recordings (R1, R2, and R3) across participants would have been desirable, it was not chosen given that negligible learning effects in repeated SRT testing conditions, only within a few tenths of a decibel, were expected [11,12]. The entire testing procedure took approximately 105 minutes, and participants were allowed to have restroom and water breaks between tests (e.g., pure-tone and SRT testing, R1 and R2 testing).

Procedural version A	Procedural version B
(n=20, 12 monolinguals and 8 bilinguals)	(n=20, 8 monolinguals and 12 bilinguals)
Familiarization	R1
R1	Familiarization
R2	R2
R3	R3

Table 3: An outline of the two procedural versions.

The level of agreement between the PTA and SRT outcomes was assessed using the Pearson product-moment correlation. A moderate positive relationship between the two measures was found, with a range between 0.411 and 0.707, and the mean coefficient of 0.558 [28].

Statistical treatment

Prior to any data analyses, the entire SRT dataset, consisting of 18 repetitions, was reviewed using an exploratory technique to identify relatively homogeneous or dissimilar groups of repetitions. The use of hierarchical cluster analysis, as a data-driven partitioning approach [29,30] was deemed appropriate given that no conclusive information is available in the literature as to what extent participants need to be familiarized in order to saturate his/her learning in repeated testing. Results from a hierarchical cluster analysis showed that only the first repetition of 18 received a different cluster membership compared to the rest (Figure 1), indicating that the first repetition was the most dissimilar from the rest. Thus, the effect of prior familiarization on the SRT outcome was tested only based on the first repetition, presumably before prior familiarization-induced SRT differences subsided. Specifically, a two-way analysis of variance (ANOVA) was performed

to examine the effects of prior familiarization (familiarized versus unfamiliarized) and language status (monolingual versus bilingual) on the SRT scores. The effects of equal syllable stress (equal stress versus unequal stress) and language status on the SRT scores were examined using a mixed model analysis with repeated measures. This analysis was conducted on the SRT data from R1 and R2 acquired from the participants who received prior familiarization, only with the exception that the first SRT repetition from R1, identified as the most dissimilar from the rest, was excluded. Another mixed model analysis with repeated measures was employed to examine the effects of the release status of the word-final stop consonants (released versus unreleased) and language status on the SRT scores. This analysis was conducted solely on the SRT data from R2 and R3 since R2 and R3 differed only in the release status of the word-final stop consonants. All statistical tests were conducted using SPSS Statistics 22.0 (IBM Corporation, Armonk, NY) at the significance level of $\alpha < 0.05$.

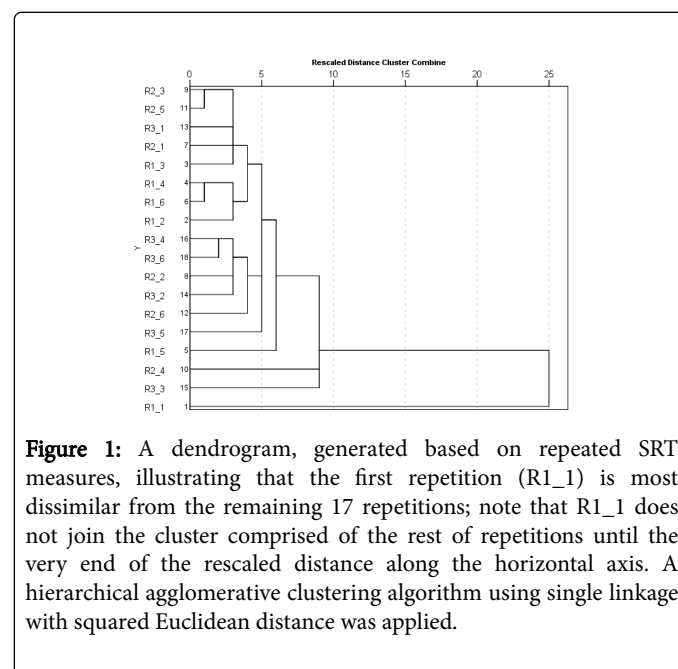


Figure 1: A dendrogram, generated based on repeated SRT measures, illustrating that the first repetition (R1_1) is most dissimilar from the remaining 17 repetitions; note that R1_1 does not join the cluster comprised of the rest of repetitions until the very end of the rescaled distance along the horizontal axis. A hierarchical agglomerative clustering algorithm using single linkage with squared Euclidean distance was applied.

Results

The participants' PTA and mean SRTs acquired from three recording conditions (R1, R2, and R3) are summarized in Figure 2. The effect of prior familiarization on the SRT outcome was tested on the first repetition of R1, while taking the participants' language status into account. Figure 3 displays SRT summary data acquired from participants with and without prior familiarization, separately computed for the monolingual and the bilingual groups. Results from the two-way ANOVA analysis showed that a statistically significant main effect was found for the prior familiarization factor ($F(1, 36)=15.306$; $p < 0.05$). The SRT mean for the group with prior familiarization (mean: 6.69 dB) was significantly lower than that for the group without prior familiarization (mean: 11.31 dB) in response to the spondees presented in R1 (unequal syllable stress with unreleased word-final stop consonants). Although the SRT mean for the monolingual group (mean: 8.19 dB) was lower than the bilingual group (mean: 9.81 dB), the difference did not reach statistical significance. In addition, the effect of prior familiarization on SRT appeared to be greater in the monolingual group than in the bilingual

group; however, the interaction between the prior familiarization and language status factors was not statistically significant.

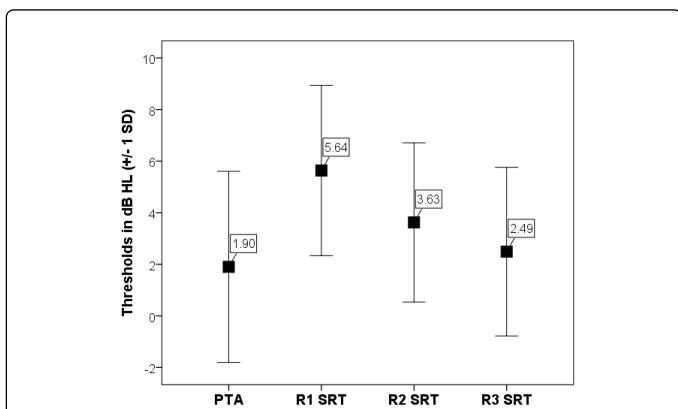


Figure 2: Means with error bars indicating +/- 1SD of the participants' PTA and SRTs from three different recording conditions: R1, spondees with unequal syllable stress and unreleased word-final stop consonants, R2, spondees with equal syllable stress and unreleased word-final stop consonants, and R3, spondees with equal syllable stress and released word-final stop consonants.

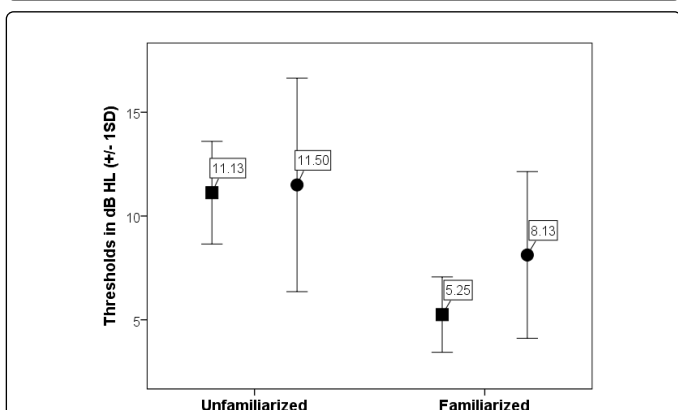


Figure 3: The SRT means acquired from two participant groups with prior familiarization (Familiarized) and without prior familiarization (Unfamiliarized), separately computed for the monolingual (square) and the bilingual (circle) participants (n=40).

The summary SRT data in Figure 4 are based on SRT outcomes from R1 and R2 (with the exception of the first repetition of R1) among participants who received prior familiarization. Results showed that the equal syllable stress factor had a statistically significant main effect on the SRT scores ($F(1, 18)=14.941$; $p<0.05$). Particularly, the participants demonstrated lower SRT scores in response to the spondees with equal syllable stress (mean: 2.92 dB) than to those with unequal syllable stress (mean: 3.99 dB). A statistically significant main effect was also found for the language status factor on SRTs ($F(1, 18)=6.121$; $p<0.05$), where the monolingual group (mean: 2.13 dB) performed better than the bilingual group (mean: 4.78 dB). No significant interaction between the syllable stress pattern and the language status factor was found.

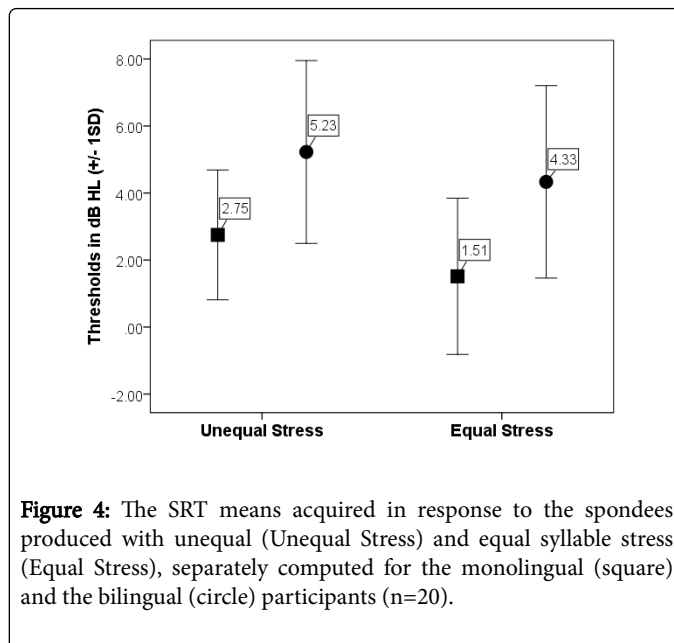


Figure 4: The SRT means acquired in response to the spondees produced with unequal (Unequal Stress) and equal syllable stress (Equal Stress), separately computed for the monolingual (square) and the bilingual (circle) participants (n=20).

Figure 5 provides summary SRT data acquired in response to the spondees with the released and unreleased word-final stop consonants for the monolingual and the bilingual participants, separately. A statistically significant main effect was found for the release status of word-final stop consonants ($F(1, 38)=26.051$; $p<0.05$). Particularly, the participants performed better in response to the spondees with released word-final stop consonants (mean: 2.49 dB) than to those with unreleased word-final stop consonants (mean: 3.63 dB). The main effect for the language status factor was also found to be statistically significant ($F(1, 38)=6.238$; $p<0.05$), with the monolingual group (mean: 1.91 dB) outperforming the bilingual group (mean: 4.21 dB). There was no significant interaction between the two main factors.

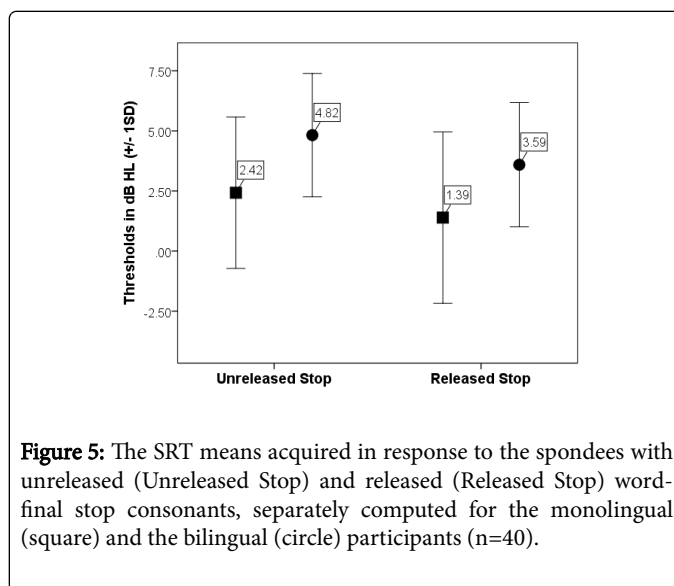


Figure 5: The SRT means acquired in response to the spondees with unreleased (Unreleased Stop) and released (Released Stop) word-final stop consonants, separately computed for the monolingual (square) and the bilingual (circle) participants (n=40).

Discussion

The purpose of this study was to identify the possible influence of familiarization, stress, and word-final release of stop consonants on

SRT outcomes in individuals with normal hearing, while taking the participants' language status (monolingual versus bilingual) into account. Not surprisingly, results showed that the group with prior familiarization had significantly better SRT scores than the group without prior familiarization. Acoustic-phonetic variations of spondee production, represented as equal versus unequal syllable stress and released versus unreleased word-final stop consonants, were found to have small but statistically significant effects on the participants' SRT outcomes. The monolingual listeners consistently outperformed their bilingual counterparts, with the exception that both groups performed poorly when no prior familiarization was provided, with a minimal between-groups difference.

As the ASHA policy document illustrates, familiarization with test vocabulary in SRT testing helps to eliminate the effects of a patient's prior knowledge of test vocabulary on the SRT outcome [2]. By the same token, familiarization lessens the degree to which patients are penalized in SRT testing due to a lack of prior knowledge of test vocabulary. Results from the present study showed that the group with prior familiarization outperformed the group without prior familiarization by 4.63 dB HL on the very first repetition of SRT testing. This difference falls within the expected range of SRT gains through prior knowledge of test vocabulary [11]. In agreement with previous studies [9,11,12], this finding suggests that prior familiarization of listeners with test vocabulary should remain an important step in SRT testing given the appreciable size of performance improvement (about 5 dB HL) that would likely decrease the SRT-PTA gap. Familiarizing patients with spondees and thus obtaining the SRT that is congruent with PTA may prevent unnecessary additional, time-consuming retesting.

The results revealed statistically significant SRT differences between the monolingual and the bilingual groups, particularly with the datasets that were used to examine the effects of equal syllable stress and of word-final stop consonant release on SRTs. The only exception was found for the testing condition in which no prior familiarization with test vocabulary was provided. Without prior familiarization, both groups performed poorly on SRT testing regardless of language status, which might have significantly mitigated the effect of the participants' language status on the SRT outcome in this particular analysis. Although it appeared that the SRT gains through prior familiarization were greater in the monolingual group than the bilingual group, no significant interaction between prior familiarization and language status was observed.

Previous research has shown that bilinguals may perform speech recognition tasks as equally well as their monolingual peers in quiet; degraded listening conditions, such as noise or reverberation, however, have been found to more adversely affect speech recognition performance in bilinguals than monolinguals despite normal auditory thresholds [31-34]. The current finding that monolinguals overall outperform bilinguals in the absence of noise has not been as widely reported in the literature. It is possible that the low decibel level employed during SRT testing, which is meant to be the softest level at which listeners can understand the presented stimuli, might have had an effect similar to noise or reverberation on bilingual participants. Nonetheless, it should also be acknowledged that the statistically significant differences between the monolingual and the bilingual groups remain less than 3 dB HL. Recall that the participants in this study identified themselves as having native-like English proficiency. Future research could investigate the replicability of the current finding, while controlling for other variables important in bilingual

research; for example, von Hapsburg and Peña [35] identified participant-related key factors of bilingualism, including language status, history, stability, competency, and demand. Additionally, a larger sample size in future studies may render more robust results in relation to the effect of language status and its interactions with common procedural variations in SRT testing.

Unlike prior familiarization, which has long been regarded as having a clinical effect on the SRT outcome, very few attempts have been made to examine acoustic-phonetic variations of spondee production and their influence on SRT outcomes. No previous study examined the effects of altering syllable stress properties on SRT outcomes, despite the likelihood that clinicians produce unequal syllable stress in spondees (with respect to f_0 and duration) [13]. Similarly, release status of the word-final stop consonants had only a theoretical basis for its potential influence on SRT outcomes [19,20]. Results from the present study showed that the participants with normal hearing demonstrated improved SRT performance in response to the spondee stimuli produced with equal syllable stress (1.1 dB HL improvement) and released word-final stop consonants (1.13 dB HL improvement). Although statistically significant, these measurable SRT changes in response to acoustic-phonetic alterations of spondee production were not as large as expected. Rather, these SRT differences seem to fall within the range of variation in repeated testing as reported in previous studies. Among the few studies that examined speech reception in repeated testing with repeated listening stimuli, minimal SRT improvement was reported with the range of 0.3-0.4 dB across repeated trials or sessions [11,12]. Therefore, the relatively small magnitude of SRT change, especially in the repeated SRT testing condition, may be interpreted as acoustic-phonetic variations in spondee production likely having little influence on the clinical interpretation of the SRT outcomes, especially in individuals with normal hearing. This conclusion may be further safeguarded by the relatively large tolerance range of SRT-PTA discrepancy; agreement between the SRT and PTA within 6 dB is generally considered normal, where the SRT tends to be higher than the PTA [36,37].

While limited practical and clinical importance is suggested in regard to controlling acoustic-phonetic variations of spondee production in SRT testing for individuals with normal hearing, some caution should be exercised when extending the current findings to the audiology clinic. Previous studies have demonstrated that stressed syllables facilitate faster phoneme and word recognition than unstressed ones; likewise, incorrect stress patterns are known to delay word recognition [17,38-41]. Given the acoustic salience of the stressed syllables, it is reasonable to hypothesize that primary stress on both syllables rather than primary stress on either syllable alone will more likely assist individuals with hearing loss in identifying the target spondee. Word-final stop release and its theoretical benefits can also be discussed in relation to clear speech. Clear speech is known to substantially enhance intelligibility, benefiting various listener populations, and one of the clear speech modifications includes increased frequency of word-final stop release [42,43]. A general hypothesis can be constructed that released word-final stop consonants will be more likely to ease word recognition tasks than their unreleased counterparts, due to additional acoustic cues provided. A possible exception, however, is that patients with high-frequency hearing loss may find this feature not as helpful, given that voiceless stop bursts (e.g., /t/ or /k/) have the noise spectrum highly concentrated in high frequency ranges [44-46]. Although the ASHA guidelines [3] do not promote or advise against the release of word-final stops, the current

study raises attention to careful administration of SRT testing in clinical practice, especially in regards to spondee production.

It may be unreasonable to expect clinicians to utilize reference tones and metronomes during monitored live voice SRT testing in order to maintain more equal f_0 s and durations between syllables. Nonetheless, the methodological approach in the creation of the recorded test stimuli as demonstrated in the present study could be utilized during production of professional recordings. This inclusion could aid in production of test material, which would more closely exemplify the definitional and theoretical spondaic requirement of equal stress. Likewise, ensuring release of word-final stop consonants could also be easily incorporated into practice during monitored live-voice testing or professional recordings. Careful attention paid toward acoustic-phonetic variations of spondee production may help clinicians provide consistency in presenting test stimuli within and across patients.

A note regarding the importance of utilizing pre-recorded spondee materials is also worthwhile. The present study utilized a standard laptop computer to administer the test material through a calibrated audiometer. The use of digitally recorded audio files via software media players can now be easily utilized and controlled as monitored live-voice testing. Such a method is simple to calibrate, administer, and offers more reliable and, arguably, more valid presentation of the test stimuli. With the ubiquity of laptops, tablets, and mp3 players, clinicians may indeed bring ASHA's guidelines [2] for obtaining the SRT in line with current technological trends.

While control of syllable duration, f_0 , and release status of final stops were given close attention, deviation in the production of these variables was inevitable given the organic nature of the human speech mechanism. Particularly, duration of individual vowel and consonantal segments, vocal quality of the speaker, and overall enunciation of the spondees were subject to variation. Such extraneous variables were inherent limitations to the present study. With the technological advent of humanlike synthetic speech, future research utilizing synthetic speech stimuli, which differ only in certain variables of interest and remain identical in all other regards, may be a viable option for ideal recordings.

Conclusion

The present study examined the individual effects of variables for obtaining SRT outcomes, which relate to commonly reported and likely procedural variation within the ASHA guidelines [2] for obtaining SRT outcomes. In agreement with previous studies [11,12], results from the present study demonstrated that prior familiarization with test vocabulary significantly improved the participants' SRT outcomes, lending strong support to the current ASHA guideline. It was also found that the effects of equal syllable stress and word-final stop release on SRTs were statistically significant. However, the observed magnitude of change, slightly greater than 1 dB HL, suggests that control of acoustic-phonetic variations of spondee production may have limited clinical importance when SRT testing is administered in individuals with normal hearing. In other words, leniency in controlling the syllable stress and word-final stop release patterns of spondee production may thus be tolerated and have little clinical influence on SRT outcomes. Nonetheless, careful and consistent presentation of test stimuli is necessary to avoid any unwanted influence from procedural variation. Future investigation that addresses the extent to which these controllable procedural variations

affect SRT outcomes in individuals with hearing impairments is warranted.

Acknowledgments

This work was partly supported by the Graduate Research Enhancement Grants (GREG) from the Office of the Vice President for Research at New Mexico State University, Las Cruces, NM. We also thank Wendy Eggebraaten for her assistance in perceptual ratings of the spondee stimuli.

References

1. Martin FN, Clark JG (2009) Introduction to audiology. (10th edn.), Pearson Education, Inc.
2. American Speech-Language-Hearing Association (1988) Determining Threshold Level for Speech.
3. Fletcher H, Steinberg JC (1929) Articulation testing methods. *Bell Sys Tech J* 8: 806-854.
4. Hudgins CV, Hawkins JE, Karlin JE, Stevens SS (1947) The development of recorded auditory tests for measuring hearing loss for speech. *Laryngoscope* 57: 57-89.
5. Hirsch IJ, Davis H, Silverman SR, Reynolds G, Eldert E, et al. (1952) Development of materials for speech audiometry. *J Speech Hear Disord* 17: 321-337.
6. Wiley TL, Stoppenbach DT, Feldhake LJ, Moss KA, Thordardottir ET (1995) Audiologic practices: What is popular versus what is supported by evidence. *Am J Audiol* 4: 26-34.
7. Debow A, Green WB (2000) A survey of Canadian audiological practices: puretone and speech audiometry. *J Speech Language Pathol Audiol* 24: 153-161.
8. Martin FN, Champlin CA, Chambers JA (1998) Seventh survey of audiometric practices in the United States. *J Am Acad Audiol* 9: 95-104.
9. Conn M, Dancer J, Ventry IM (1975) A spondee list for determining speech reception threshold without prior familiarization. *J Speech Hear Disord* 40: 388-396.
10. Wilson RH, Margolis RH (1983) Measurements for auditory measurements for speech stimuli. In: Konkle DF, Rintelmann WF (edn.) Principles of speech audiometry. University Park Press, pp. 79-126.
11. Tillman TW, Jerger J (1959) Some factors affecting the threshold in normal hearing subjects. *J Speech Hear Res* 2: 141-146.
12. Jerger JF, Carhart R, Tillman TW, Peterson JL (1959) Some relations between normal hearing for pure tones and for speech. *J Speech Hear Res* 2: 126-140.
13. Bettagere R (2012) Acoustic characteristics of spondee syllable stress. *Asia Pacific J Speech Lang Hear* 15: 29-40.
14. Lieberman P (1960) Some acoustic correlates of word stress in American English. *J Acoust Soc Am* 32: 451-454.
15. Ying GS, Jamieson LH, Chen R, Michell CD (1996) Lexical stress detection on stress-minimal word pairs. Proceedings of the 1996 International Conference on Spoken Language Processing.
16. Albin DD, Echols CH (1996) Stressed and word-final syllables in infant-directed speech. *Infant Behav and Dev* 19: 401-418.
17. Cutler A, Foss DJ (1977) On the role of sentences stress in sentences processing. *Lang Speech* 20: 1-10.
18. Shriberg LD, Kent RD (2003) Clinical Phonetics (3rd edn.), Pearson Education Inc.
19. Malécot A (1959) The role of releases in the identification of released final stops; a series of tape-cutting experiments. *Lang* 34: 370-380.
20. Lisker L (1999) Perceived final voiceless stops without release: effects of preceding monophthongs versus nonmonophthongs. *Phonetica* 56: 44-55.

21. Ramkissoon I, Proctor A, Lansing CR, Bilger RC (2002) Digit speech recognition thresholds (SRT) for non-native speakers of English. *Am J Audiol* 11: 23-28.
22. Mayo LH, Florentine M, Buus S (1997) Age of second-language acquisition and perception of speech in noise. *J Speech Lang Hear Res* 40: 686-693.
23. Takata Y, Nabelek AK (1990) English consonant recognition in noise and in reverberation by Japanese and American listeners. *J Acoust Soc Am* 88: 663-666.
24. Shin HB, Kominski RA (2010). *Language Use in the United States: 2007 American Community Survey Reports*. US Census Bureau.
25. Zlatin MA, Koenigsnecht RA (1976) Development of the voicing contrast: a comparison of voice onset time in stop perception and production. *J Speech Hear Res* 19: 93-111.
26. Colton RH, Casper JK, Leonard R (2011) *Understanding voice problems: A physiological perspective for diagnosis and treatment* (4th edn.). Lippincott Williams & Wilkins.
27. Boersma P (2001) Praat, a system for doing phonetics by computer. *Glott International* 5: 341-345.
28. Dancey C, Reidy J (2004) *Statistics without maths for psychology: Using SPSS for Windows*. Prentice Hall, London.
29. Berkhin P (2006) A survey of clustering data mining techniques. In: Kogan J, Nicholas C, Tebouille M (edn.) *Grouping Multidimensional Data*. Springer, pp. 25-72.
30. Loureiro A, Torgo L, Soares C (2004) Outlier detection using clustering methods: a data cleaning application. *Proceedings of KDNNet Symposium on Knowledge-Based Systems for the Public Sector*, Bonn, Germany.
31. Florentine M (1985) Speech perception in noise by fluent, non-native listeners. *J Acoust Soc Am* 77: S106.
32. Nabelek A, Donahue A (1984) Perception of consonants in reverberation by native and non-native listeners. *J Acoust Soc Am* 75: 632-634.
33. Rogers CL, Lister J, Febo DM, Besing JM, Abrams JB (2006) Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Appl Psycholinguist* 27: 465-485.
34. Von Hapsburg D, Champlin CA, Shetty SR (2004) Reception thresholds for sentences in bilingual (Spanish/English) and monolingual (English) listeners. *J Am Acad Audiol* 15: 88-98.
35. Von Hapsburg D, Peña E (2002) Understanding bilingualism and its impact on speech audiometry. *J Speech Lang, Hear Res* 45: 202-213.
36. Lawson G, Peterson M (2011) *Speech Audiometry*. Plural Publishing.
37. Letowski T, Hergenreder P, Houchin T (1992) Relationships between speech recognition threshold, average hearing level, and speech importance noise detection threshold. *J Speech Hear Res* 15: 1131-1136.
38. Connine CM, Clifton C Jr, Cutler A (1987) Effects of lexical stress on phonetic categorization. *Phonetica* 44: 133-146.
39. Cutler A, Clifton C Jr (1984) The use of prosodic information in word recognition. In: Bouma H, Bouwhuis DG (edn.) *Attention and Performance X*. Erlbaum, pp. 183-196.
40. Slowiaczek LM (1990) Effects of lexical stress in auditory word recognition. *Lang Speech* 33: 47-68.
41. Slowiaczek LM, Soltano EG, Bernstein HL (2006) Lexical and metrical stress in word recognition: lexical or pre-lexical influences? *J Psycholinguist Res* 35: 491-512.
42. Picheney MA, Durlach, NI, Braida LD (1986) Speaking clearly for the hard of hearing. II: Acoustic characteristics of clear and conversational speech. *J Speech Hear Res* 29: 434-446.
43. Bradlow AR, Kraus N, Hayes E (2003) Speaking clearly for children with learning disabilities: sentence perception in noise. *J Speech Lang Hear Res* 46: 80-97.
44. Halle M, Hughes GW, Radley JPA (1957) Acoustic properties of stop consonants. *J Acoust Soc Am* 29: 107-116.
45. Zue VW (1976) *Acoustic characteristics of stop consonants: a controlled study* (Tech Report No. 523). Lincoln Laboratory, MIT.
46. Parikh G, Loizou P (2005) The influence of noise on vowel and consonant cues. *J Acoust Soc Am* 118: 3874-3888.