J Am Acad Audiol 13: 154–159 (2002)

Effects of Stimulus Presentation Level on Stop Consonant Identification in Normal-Hearing and Hearing-Impaired Listeners

Patrick N. Plyler*
Mark S. Hedrick†

Abstract

The purpose of the present study was to determine whether varying the presentation level of stop consonant stimuli resulted in similar phonetic boundary shifts for listeners with normal and impaired hearing. Sixteen normal-hearing and 16 hearing-impaired listeners categorized synthetic speech stimuli as /b/, /d/, or /g/. The onset frequency of F2 varied from 900 to 2300 Hz (100-Hz steps), and the presentation level varied from 92 to 62 dB SPL (10-dB steps) for each stimulus presentation. Hearing-impaired listeners had significantly more missing boundary values than normal-hearing listeners; however, the correlation between the number of missing boundary values and hearing sensitivity was not significant. Comparison of boundary shift with level demonstrated that hearing-impaired listeners had a smaller boundary shift with increasing level than normal-hearing listeners. The amount of boundary shift was not correlated with audibility. The results of the current study suggest that increasing the presentation level of a signal does not result in performance similar to that of listeners with normal hearing.

Key Words: Hearing loss, phonetic boundary, speech perception, stop consonant identification

Abbreviations: CSRE = Computerized Speech Research Environment; FFR = frequency following response

Sumario

El propósito del presente estudio fue determinar si la variación del nivel de presentación de estímulos representados por consonantes oclusivas produce cambios similares de límite fonético para sujetos con audición normal y alterada. Dieciséis sujetos normo-oyentes y dieciséis hipoacúsicos evaluaron estímulos sintéticos de lenguaje, tales como /b/, /d/, o /g/. La frecuencia F2 de inicio varió desde 900 a 2300 Hz (pasos de 100 Hz), y el nivel de presentación varió desde 92 a 62 dB SPL (pasos de 10 dB) para cada estímulo presentado. Los sujetos hipoacúsicos mostraron alteración de los valores limítrofes más significativamente que los sujetos normo-oyentes, sin embargo, la correlación entre los valores limítrofes alterados y la sensibilidad auditiva no fue significativa. La comparación del cambio de límite fonético con el nivel de presentación demostró que los sujetos hipoacúsicos tenían un cambio de límite más pequeño conforme se incrementaba la intensidad, que los sujetos normo-oyentes. La magnitud de este cambio de límite fonético no correlacionó con la audibilidad. Los resultados de este estudio sugieren que el incremento en el nivel de presentación de una señal no mejora el rendimiento en la forma en que ocurre con sujetos normo-oyentes.

Palabras Clave: Pérdida auditiva, límite fonético, percepción del lenguaje, identificación de consonantes oclusivas

Abreviaturas: CSRE = Ambiente de Investigación de Lenguaje Computarizado; FRR = respuesta de seguimiento frecuencial

*Department of Communication Sciences and Disorders, Louisiana State University, Baton Rouge, Louisiana; †Department of Audiology and Speech Pathology, University of Tennessee, Knoxville, Tennessee

Reprint requests: Patrick N. Plyler, Department of Communication Sciences and Disorders, Louisiana State University, 163 Music and Dramatic Arts Building, Baton Rouge, LA 70803

Listeners with normal hearing are able to categorize stop consonant place of articulation when the onset frequency of the second-formant transition is varied (Liberman et al, 1954; Dorman and Dougherty, 1981; Kewley-Port, 1981). Listeners with mild to moderate sensorineural hearing loss, however, have difficulty using the second-formant transition for the identification of stop consonants (Dorman et al, 1985). For example, a study investigating the effects of varying the onset frequency of the second-formant transition on the identification of stop consonant place of articulation in listeners with normal and impaired hearing revealed identification functions with significantly different phonetic boundaries (the point of 50% identification accuracy) for each group (Dorman et al, 1985). Differences in phonetic boundaries suggested that listeners with cochlear damage have a degraded ability to categorize stop consonant place of articulation based solely on the dynamic spectral information present in the second-formant transition (Dorman et al, 1985).

Altering the presentation level of stop consonant stimuli shifts the phonetic boundaries of stop consonant identification functions in listeners with normal hearing (Dorman and Dougherty, 1981). As the presentation level of the stimulus decreases, the phonetic boundary increases in frequency. Therefore, decreasing the presentation level of the stimulus results in an enlargement of the /b/ identification function, a shifting or reduction of the /d/ identification function, and a narrowing of the /g/ identification function in listeners with normal hearing (Dorman and Dougherty, 1981). It is unclear, however, whether altering the presentation level of stop consonant stimuli results in phonetic boundary shifts in listeners with impaired hearing. Such a manipulation may reveal differences in the perception of stops caused by a sensorineural hearing loss and may prove important since individuals with hearing loss need higher presentation levels to overcome loss of sensitivity. Thus, the purpose of the present study was to determine whether altering the presentation level of stop consonant stimuli resulted in similar phonetic boundary shifts for listeners with normal and impaired hearing.

METHOD

Subjects

Thirty-two subjects participated in this experiment. The subjects were divided into a normal-hearing group (22 to 40 years in age) and a hearing-impaired group (20 to 67 years in age). The criteria for inclusion in the normal-hearing group included (1) hearing sensitivity of 20 dB HL (ANSI, 1996) or better for octave frequencies from 250 to 8000 Hz in the right ear, (2) normal appearance of ear canal and pinna, (3) normal tympanograms bilaterally, and (4) no air–bone gaps greater than 10 dB. The inclusion criteria for the hearing-impaired group differed from that of the normal-hearing group in one aspect: hearing sensitivity between 20 and 60 dB HL for octave frequencies from 250 to 2000 Hz in the right ear (Fig. 1).

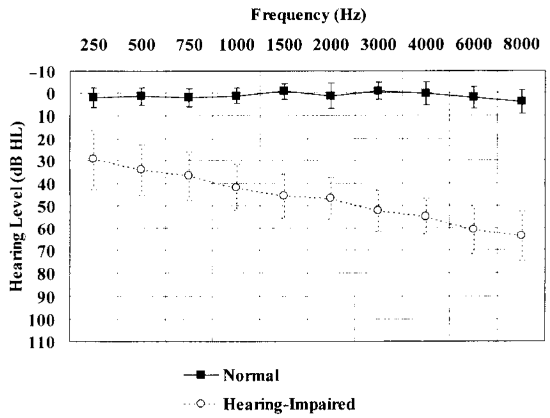

Figure 1 Mean air-conduction thresholds and standard deviations of the test ear of the normal-hearing and hearing-impaired subjects.

Stimuli

A 15-step /bα/-/dα/-/gα/ continuum was generated using the cascade configuration of Klatt's formant synthesizer (Klatt, 1980) at a sampling rate of 10 kHz. The total duration of each stimulus was 100 msec, whereas the duration of the first- and second-formant transitions was 40 msec. The starting frequency of the first formant was 150 Hz, which increased until it reached steady state at 750 Hz. The onset frequency of the second-formant varied from 900 to 2300 Hz in 100-Hz steps to create the /bα/-/dα/-/gα/ perceptual continuum. The frequencies of the third and fourth formants were 2400 and 3300 Hz throughout the duration of the stimulus (Fig. 2). The amplitude of the second-formant was –3 dB relative to the first formant; the amplitude of the third formant was –6 dB relative to the first formant.

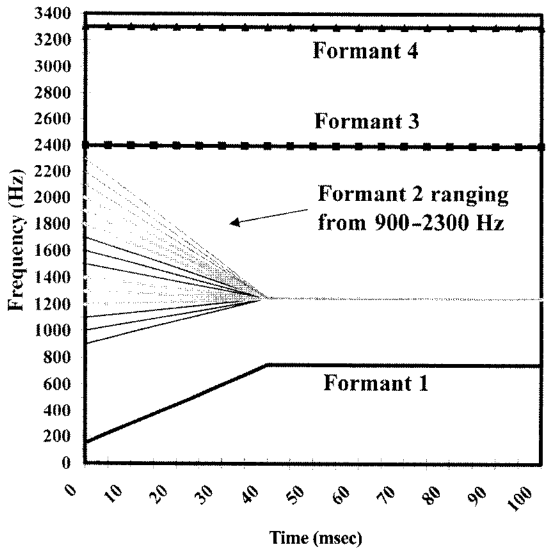

Figure 2 Schematic of the synthetic speech stimuli.

Apparatus

Stop consonant identification was evaluated using the 15 stimuli and four presentation levels (92, 82, 72, 62 dB SPL). Four levels were used to determine if a systematic boundary shift was present. The stop consonant stimuli were digitally controlled by a signal generation system (Tucker-Davis Technologies, System II) interfaced to a microcomputer (Compaq 2000, 166 MHz). Digital signal generation, including control of parameters such as overall intensity and programming of the research protocol, was accomplished by interactive signal generation and control software (Computerized Speech Research Environment [CSRE], Version 4.5). The digital stimuli were routed from the microcomputer to the digital-to-analog converter (TDT-DD1). The output of the converter was routed to a low-pass filter (TDT-PF1, 4.9 kHz low pass), then to a programmable attenuator (TDT-PA4), and then to the headphone buffer (TDT-HB) before being delivered to an insert earphone (Etymotic, ER-3A) located in a sound-treated examination room (Industrial Acoustics Company).

Procedures

Earphone output was measured and calibrated to peak vowel amplitude levels of 62, 72, 82, and 92 dB SPL prior to each test session. Subjects were seated in the examination room and were instructed to view a computer monitor through the examination room window. Following each stimulus presentation, subjects indicated which stop consonant they perceived by selecting the appropriate symbol displayed on the monitor (B, D, or G) via the computer mouse. Each subject participated in a trial session prior to experimental testing. An initial trial session consisted of 10 tokens of each of the stimuli judged by two independent listeners to represent the best exemplars of /bɑ/, /dɑ/, and /gɑ/. Testing commenced when it had been established that the subject could correctly identify 70 percent of the trial items presented at 82 dB SPL.

In the experimental session(s), each of the 15 stimuli was randomly presented at each of the four levels over 20 trials for a total of 1200 responses. Generation of random stimulus orders and online data collection were performed using a commercially available identification program (CSRE, Version 4.5). Psychometric functions displaying the percentage identification as a function of onset frequency were generated for each stop consonant at each level for each subject and group.

RESULTS

Boundary locations corresponding to 50 percentage points were calculated in Hertz from each listener's psychometric functions. Four boundary values were computed: one corresponding to /b/, one for the /d/ category adjacent to /b/, another for the /d/ category adjacent to /g/, and one for the /g/ category.

From the data collection phase, it was obvious that some listeners had psychometric functions that did not yield a straightforward boundary calculation (6 normal-hearing and 10 hearing-impaired listeners). Further examination of the boundary data per listener revealed a Poisson distribution, with 16 listeners having no missing boundary values, 3 listeners having 2 missing values, and so on until there was a single listener with all 16 boundary values missing. The results of a one-sided nonparametric Wilcoxon test verified that the normal-hearing listeners had significantly fewer missing values (mean = 1.3) compared to the hearing-impaired listeners (mean = 4.3, $W = 221.5$, $p = .049$). To examine the relationship between the number of missing values and the audiometric pure-tone average at 1.0, 1.5, and 2.0 kHz, a Spearman correlation was computed. It was not significant ($p = .187$).

Phonetic boundary shift was calculated for all subjects with no missing boundary values (10 normal-hearing and 6 hearing-impaired listeners). Boundary shift was determined by sub-

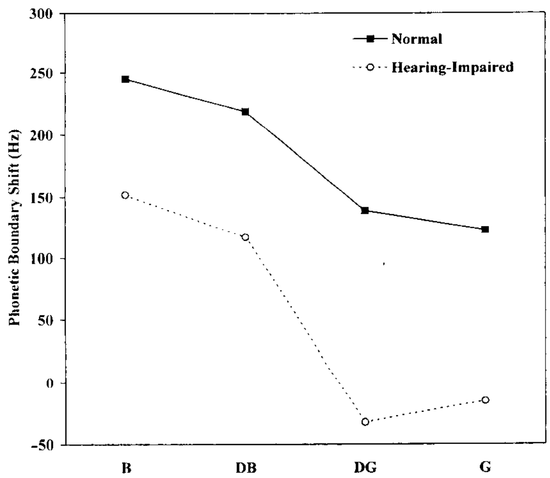

Figure 3 Mean phonetic boundary shift for each group and for each category. Values reflect the amount of phonetic boundary shift (in Hz) when the stimulus level decreased from 92 to 62 dB SPL for subjects with no missing values.

tracting the boundary value obtained at 92 dB SPL from the boundary value obtained at 62 dB SPL to measure the effect of decreasing signal level. Mean boundary shift data for subjects with no missing boundary values are displayed in Figure 3.

The boundary shift values were used as the dependent variable in a repeated-measures analysis of variance to determine if the amount of boundary shift was similar for each group and for each phoneme (Table 1). The analysis revealed that the amount of boundary shift was significantly greater for the normal-hearing group than for the hearing-impaired group ($F = 5.05$, $df = 1, 14$, $p < .05$). Furthermore, the analysis also revealed a significant main effect for phonemic category ($F = 3.64$, $df = 3, 42$, $p < .05$). The results of the paired samples t-tests indicated that the boundary shift of the /b/ category was significantly greater than that of the /d/ adjacent to /b/ (DB), /d/ adjacent to /g/ (DG), or /g/ categories (Table 2). The group-by-phoneme interaction was not significant ($F = 0.20$, $df = 3, 42$, $p > .05$). To examine the relationship between the amount of boundary shift and the audiometric pure-tone average at 1.0, 1.5, and 2.0 kHz, a Pearson correlation was computed for each category. No comparisons were significant (Table 3).

Table 1 Results of Between-Subjects Analysis of Variance Completed on the Boundary Shift Data

Effect	*df*	*F value*	*p*
Group	1, 14	5.05	.041
Phoneme	3, 42	3.64	.020
Phoneme × Group	3, 42	0.20	.894

Table 2 Results of Paired-Samples t-Tests on the Boundary Shift Data

Effect	*df*	*t Value*	*p*
B–DB	15	−2.36	.032
B–DG	15	−2.12	.050
B–G	15	−2.21	.043
DB–DG	15	−1.60	.130
DB–G	15	−1.67	.114
DG–G	15	−0.43	.670

DISCUSSION

The results of the present study indicate that hearing-impaired listeners demonstrate greater difficulty than normal-hearing listeners in categorizing stop consonant place of articulation as evidenced by the hearing-impaired group's significantly greater number of missing boundary values. Furthermore, the results suggest that varying the presentation level of the signal resulted in less phonetic boundary shift for listeners with impaired hearing than for listeners with normal hearing.

One possible explanation for the degraded performance exhibited by the hearing-impaired group may be related to the fact that the static burst was not included in the synthesis of the stimulus. Listeners with impaired hearing who are unable to base phoneme identification on the dynamic spectral information of the speech waveform may place greater emphasis on the static spectral information in the waveform to accurately discriminate phoneme place of articulation. Therefore, hearing-impaired listeners in

Table 3 Results of Pearson Correlation of Boundary Shift Data and Audiometric Pure-Tone Average at 1.0, 1.5, and 2.0 kHz

Effect	*Pearson Correlation*	*p*
B	.352	.182
DB	.383	.143
DG	.286	.282
G	.203	.451

the present study who normally rely on static cues would have demonstrated more difficulty than listeners who normally rely on dynamic cues. This reasoning could also explain the small number of missing boundary values in the normal-hearing group. Some listeners with normal hearing may also rely more heavily on burst information than transitional information (Hazan and Rosen, 1991). Therefore, some normal-hearing listeners would have difficulty with the categorization task. Thus, it is plausible to reason that the absence of static information could have contributed to group performance differences.

A second possible explanation for the degraded performance exhibited by the impaired group may be related to the audibility of the second-formant transition. For example, the mean audiometric pure-tone average at 1.0, 1.5, and 2.0 kHz of the hearing-impaired group was 44 dB greater than that of the normal-hearing group. Thus, the audibility of the second-formant transition was reduced for the hearing-impaired listeners. This second explanation, however, is not supported by the results of the current study. Had decreased audibility of the second-formant transition caused listeners with impaired hearing to have a greater number of missing boundary values, a relationship would have existed between the number of missing values and the hearing sensitivity of the listener. Stated differently, listeners with the most hearing impairment would have the highest number of missing boundary values, whereas listeners with the least hearing impairment would have the lowest number of missing boundary values. However, the results of the present study indicate that the number of missing boundary values is not correlated with the hearing sensitivity of the listener. Listeners with near-normal pure-tone averages (1.0, 1.5, and 2.0 kHz) for whom audibility was improved were as likely to have indiscernible boundaries as were listeners with greater degrees of hearing loss for whom audibility was reduced. Given this, it is unlikely that the identification difficulties evident in the impaired group were solely attributable to audibility differences between the groups.

Likewise, had decreased audibility of the second-formant transition caused listeners with impaired hearing to have a reduced phonetic boundary shift, a relationship would have existed between the amount of phonetic boundary shift and the hearing sensitivity of the listener. Stated differently, listeners with the most hearing impairment would have less phonetic boundary shift, whereas listeners with the least hearing impairment would have the most phonetic boundary shift. However, the results of the present study indicate that the amount of phonetic boundary shift was not correlated with the hearing sensitivity of the listener. Given this, it is unlikely that the reduced phonetic boundary shifts evident in the hearing-impaired group were solely attributable to audibility differences between the groups.

Performance difficulties evident in the hearing-impaired group may have been in part caused by a degraded ability to use dynamic spectral information present in the second-formant transition of the stimulus. Auditory nerve single-unit studies have demonstrated that phase-locking plays an important role in the neural encoding of the spectrum of speech-like sounds. For example, neurophysiologic studies performed on cats have indeed shown degradation of the neural representation of speech-like sounds following acoustic trauma (Miller et al, 1997). Specifically, following acoustic trauma, fibers in the damaged region of the cochlea respond to components of the stimulus over a broad frequency range. Thus, the responses are no longer frequency specific. Miller et al suggested that human listeners with sensorineural hearing loss would have problems with the identification of speech spectral peaks. Formant transitions represent speech spectral peaks that are changing over time. Thus, there is possibly a neurophysiologic basis for listeners with sensorineural hearing loss to have disrupted encoding of formant transition cues.

Recently, it has been reported that the phase-locked activity underlying the scalp-recorded human frequency following response (FFR) encodes the first two formants of several steady-state vowels (Ananthanarayan, 1999), and the time-variant frequency presented in tonal sweeps (Ananthanarayan and Parkinson, 2000). Plyler and Ananthanarayan (2001) recorded FFRs using synthetic stop consonant stimuli in normal-hearing and hearing-impaired listeners. The results indicated that the FFR encoded the time-varying frequency of the second-formant transition in normal listeners; however, FFR encoding of the second-formant transition was severely degraded in listeners with mild to moderately severe sensorineural hearing loss (Plyler and Ananthanarayan, 2001). Furthermore, comparison of perceptual and electrophysiologic data for individual hearing-impaired listeners appeared to suggest that degradation in the neural representation of the formant transition may be accompanied by a reduction in identification performance (Plyler and Ananthanarayan, 2001). Given this, it is reasonable to postulate that identification difficulties evident in the hearing-impaired group were in part caused by a degraded ability to use dynamic spectral information present in the second-formant transition of the stimulus.

CONCLUSIONS

The results of the present study indicate that impaired listeners have difficulty using the second-formant transition for the identification of stop consonants. Furthermore, varying the presentation level of stop consonant stimuli resulted in less phonetic boundary shift in hearing-impaired listeners than in normal-hearing listeners. Correlation analysis did not reveal a significant relationship between identification performance and hearing sensitivity, suggesting that difficulties evident in the hearing-impaired group were not solely attributable to audibility differences between the groups. Lastly, single-unit and ensemble response studies suggest that performance difficulties evident in the hearing-impaired group may in part be caused by a degraded ability to use dynamic spectral information present in the second-formant transition of the stimulus.

CLINICAL IMPLICATIONS

The results of the present study indicate that increasing the audibility of the signal for hearing-impaired listeners did not result in improved performance. These findings may help explain why some hearing-impaired listeners do not respond positively to amplification. For some hearing-impaired listeners, increasing audibility through the use of a hearing instrument results in significant improvements in speech intelligibility. In the present study, however, increasing audibility did not necessarily result in improved performance. Therefore, increasing the audibility of the signal may or may not result in improved speech intelligibility for all hearing-impaired listeners.

As previously mentioned, single-unit and ensemble response studies suggest that performance difficulties evident in hearing-impaired listeners may in part be caused by a degraded ability to use dynamic spectral information present in the stimulus. Although current hearing instrument fitting methods provide valuable information to the clinician regarding the signal processing capabilities of the amplification device, it is unlikely that acoustic measurements obtained at the tympanic membrane of the listener reflect the neural representation of these stimuli following processing by the impaired auditory system in question. Given this, future hearing instrument fitting methods should attempt to determine the neural representation of the stimulus after cochlear processing rather than determine the acoustic representation of the stimulus prior to cochlear processing.

REFERENCES

American National Standards Institute. (1996). *Specification for Audiometers.* (ANSI S3.6-1996). New York: ANSI.

Ananthanarayan AK. (1999). Human frequency-following responses to two-tone approximations of steady state vowels. *J Audiol Neurootol* 4:95–103.

Ananthanarayan AK, Parkinson J. (2000). Human frequency-following responses: representation of tonal sweeps. *J Audiol Neurootol* 5:312–321.

Dorman MF, Dougherty K. (1981). Shifts in phonetic identification with changes in signal presentation level. *J Acoust Soc Am* 69:1439–1440.

Dorman MF, Marton K, Hannley MT, Lindholm JM. (1985). Phonetic identification by elderly normal and hearing-impaired listeners. *J Acoust Soc Am* 77:664–670.

Hazan V, Rosen S. (1991). Individual variability in the perception of cues to place contrasts in initial stops. *Percept Psychophys* 49:187–200.

Kewley-Port D. (1981). *Representations of Spectral Change as Cues to Place of Articulation in Stop Consonants.* PhD dissertation, Indiana University, Bloomington, Indiana.

Klatt DH. (1980). Software for a cascade/parallel formant synthesizer. *J Acoust Soc Am* 67:971–995.

Liberman AM, Delattre PC, Cooper FS, Gerstman LJ. (1954). The role of consonant-vowel transitions in the perception of the stop and nasal consonants. *Psychol Monogr* 68:1–13.

Miller RL, Schilling JR, Franck KR, Young ED. (1997). Effects of acoustic trauma on the representation of the vowel /e/ in cat auditory nerve fibers. *J Acoust Soc Am* 101:3602–3616.

Plyler PN, Ananthanarayan AK. (2001). Human frequency-following responses: representation of second-formant transitions in normal-hearing and hearing-impaired listeners. *J Am Acad Audiol* 12:523–533.