ability may well be considerably less in young speakers. Research is currently underway in our laboratory to determine the extent to which variability on frequency stability measures is age related.

ACKNOWLEDGMENTS

We would like to express our thanks to Deb Fabrycki, Katie Bronson, Cathy Armstrong, Nancy Bloom, and Peg Paulus for assistance recording subjects and analyzing data for this study.

Horii, Y. (1980). "Vocal shimmer in sustained phonation," J. Speech Hear. Res. 23, 202–209.

House, A., and Fairbanks, G. (1953). "Influence of consonantal environment on secondary characteristics of vowels," J. Acoust. Soc. Am. 25, 105–113.

Korabic, E., Kuos, S., and Linville, S.E. (1986) "Analysis of pitch perturbation (jitter) using a Visi-Pitch and a microprocessor," J. Comput. Users Speech Hear. 2 (2), 86–93.

Lehiste, I., and Peterson, G. (1961). "Some basic considerations on the analysis on intonation," J. Acoust. Soc. Am. 33, 419–425.

Linville, S.E., and Fisher, H. (1985a). "Acoustic characteristics of women's voices with advancing age," J. Gerontol. 40, 324–330.

Linville, S.E., and Fisher, H. (1985b). "Acoustic characteristics of perceived versus actual vocal age in controlled phonation by adult females," J. Acoust. Soc. Am. 78, 40–48.

Linville, S.E., and Korabic, E. (1985). "Fundamental frequency stability characteristics of elderly women's voices," Paper presented to the American Speech-Language-Hearing Association, Washington, DC.

Sorenson, D., and Horii, Y. (1983). "Frequency and amplitude perturbation in the voices of female speakers," J. Commun. Disord. 16, 57–61.

Stoicheff, M. (1981). "Speaking fundamental frequency characteristics of nonsmoking female adults," J. Speech Hear. Res. 24, 437–441.

Wilcox, K. (1978). "Speaker age and vowel differences of vocal jitter in sustained phonation," M.A. thesis, Purdue University, Indiana.

Wilcox, K., and Horii, Y. (1980). "Age and changes in vocal jitter," J. Gerontol. 35, 194–198.

Effects of acoustic modification on consonant recognition by elderly hearing-impaired subjects

Sandra Gordon-Salant
Department of Hearing and Speech Sciences, University of Maryland, College Park, Maryland 20742

(Received 19 May 1986; accepted for publication 24 November 1986)

In a recent study [S. Gordon-Salant, J. Acoust. Soc. Am. 80, 1599–1607 (1986)], young and elderly normal-hearing listeners demonstrated significant improvements in consonant–vowel (CV) recognition with acoustic modification of the speech signal incorporating increments in the consonant–vowel ratio (CVR). Acoustic modification of consonant duration failed to enhance performance. The present study investigated whether consonant recognition deficits of elderly hearing-impaired listeners would be reduced by these acoustic modifications, as well as by increases in speech level. Performance of elderly hearing-impaired listeners with gradually sloping and sharply sloping sensorineural hearing losses was compared to performance of elderly normal-threshold listeners (reported previously) for recognition of a variety of nonsense syllable stimuli. These stimuli included unmodified CVs, CVs with increases in CVR, CVs with increases in consonant duration, and CVs with increases in both CVR and consonant duration. Stimuli were presented at each of two speech levels with a background of noise. Results obtained from the hearing-impaired listeners agreed with those observed previously from normal-hearing listeners. Differences in performance between the three subject groups as a function of level were observed also.

PACS numbers: 43.71.Ky, 43.71.Es, 43.66.Sr

INTRODUCTION

A variety of techniques have been described in which the speech signal is transformed acoustically to improve speech intelligibility. Elderly listeners are potential users of speech enhancement techniques because of the prevalence of their speech recognition problems (Dubno *et al.*, 1984; Jokinen, 1973). One speech enhancement scheme incorporates some of the acoustic modifications found in naturally produced "clear speech," including increases in the consonant energy relative to the vowel energy (consonant–vowel ratio, or CVR) and increases in the duration of individual speech sounds (Chen, 1980; Picheny and Durlach, 1979).

A recent investigation (Gordon-Salant, 1986) evaluated the effectiveness of computer-generated increments of consonant duration and CVR on speech intelligibility for young and elderly normal-hearing subjects. The stimuli were syllables of the form *CV*, where *C* was one of 19 consonants and *V* was one of three vowels. They were presented in unmodified form, with consonant duration increased by 100%, with the CVR increased by 10 dB, and with a combination of both the duration and CVR increments. The results showed that the CVR increment and combined increment improved consonant recognition in noise, whereas the duration increment did not improve performance for either listener group. Moreover, the CVR modification (but not the combined increase modification) improved recognition of most consonant place, manner, and voicing classes, without a substantial increase in consonant confusions. Young listeners outperformed the elderly listeners in each condition, although scores of elderly listeners in the CVR and combined modifications were approximately the same as those of young listeners in the unmodified condition. Thus elderly listeners with normal hearing appear to benefit from enhancements of CVR for consonant recognition in noise.

The benefits of transforming "conversational speech" to "clear speech" by specific acoustic modification for elderly hearing-impaired listeners have not been investigated. The purpose of the present study was to determine if the same patterns of consonant recognition performance obtained with acoustic modification for normal-hearing listeners would be observed for elderly hearing-impaired listeners. Nonsense syllables were presented in unmodified and modified forms at two speech levels to determine whether the improvements associated with the acoustic modifications are consistent at higher speech levels, which may compensate for the sensitivity loss.

I. METHOD

A. Stimuli and noise

The stimuli and noise were identical to those presented in Gordon-Salant (1986). The test stimuli consisted of 19 consonants, /b,d,g,p,t,k,m,n,s,z,ʃ,θ,ð,f,v,w,j,l,r/, paired with each of three vowels, /ɑ,i,u/, in a CV format. The syllables were recorded by a male speaker, low-pass filtered (7-kHz cutoff), digitized (16-kHz rate), and scaled to equate peak level across CVs. A recording of 12-talker babble was used as a background noise.

Four stimulus sets were prepared for the experiment. Details of the preparation and acoustic characteristics of these stimuli can be found in Gordon-Salant (1986). Briefly, the four stimulus sets were: (1) *baseline* CVs, which did not incorporate any acoustic modification; (2) *duration increase* CVs, whose consonant duration was increased by 100%; (3) *CVR increase* CVs, whose consonant energy relative to the vowel energy was increased by + 10 dB; and (4) *combined increase* CVs, which incorporated both the duration increase and CVR increase modifications.

B. Procedures

Testing conditions were also identical to those used in the earlier experiment (Gordon-Salant, 1986). Specifically, the nonsense syllable stimuli were presented in three vowel contexts (/ɑ,i,u/), four acoustic modifications (baseline, duration increase, CVR increase, combined increase), and two speech levels (75 and 90 dB SPL) for a total of 24 listening conditions. Stimuli were presented at a + 6-dB signal-to-babble ratio (S/B) to the better ear of each listener via a head-mounted earphone (TDH-49). During each condition, ten randomized tokens of the 19 CVs in one vowel context and at one speech level were presented for identification. Stimulus presentation and response collection were controlled by a computerized test system. Listeners' responses were limited to 19 choices, displayed in alphabetic order and orthographic form on a response box.

C. Subjects

Two groups ($N = 10$/group) of elderly (65–75 years of age) hearing-impaired subjects participated in the experiment. The first group had gradually sloping mild-to-moderate sensorineural hearing losses, with pure-tone thresholds (PTTs) decreasing by 5–15 dB/oct between 500 and 4000 Hz, and pure-tone averages (PTAs) < 40 dB HL (ANSI, 1969). The second group had sharply sloping sensorineural hearing losses, as evidenced by a difference in PTTs between 500 through 4000 Hz > 45 dB or a difference in PTTs between adjacent octave frequencies exceeding 30 dB. In addition, data from ten elderly (65–75 years of age) subjects with PTTs within normal limits ($\leqslant 15$ dB HL, 250–8000 Hz) reported previously (Gordon-Salant, 1986) served as a normal-hearing referent with which to compare the performance of the two hearing-impaired groups.

II. RESULTS AND DISCUSSION

Figure 1 presents the mean overall percent-correct consonant recognition scores from the three subject groups for syllables presented in three vowel contexts, at two speech levels, with four acoustic modifications. The effect of acoustic modification was significant [$F(3,81) = 179.21$, $p < 0.01$], and did not interact with either presentation level or listener group. However, acoustic modification did interact significantly with vowel context [$F(6,162) = 13.17$, $p < 0.01$]. Simple main effects and multiple comparison tests revealed that in each vowel context, the CVR increase and combined increase modifications yielded higher scores than were observed in the baseline and duration increase conditions ($p < 0.01$). The source of the interaction between acoustic modification and vowel was a larger improvement with the CVR and combined increase modifications over the baseline and duration increase modification in the /i/ and /u/ contexts than in the /ɑ/ context.

These results show that elderly listeners with gradually sloping and sharply sloping hearing losses exhibit comparable patterns of recognition performance with acoustic modification as do elderly listeners with normal thresholds. The success of the CVR enhancement is probably associated with an effective increase in the signal (consonant)-to-noise ratio, which should improve speech recognition performance according to articulation index theory (French and Steinberg, 1947). The minimal effects of the duration increase modification may be associated with a lack of a processing problem for stimulus duration cues among the elderly subjects (Dorman *et al.*, 1985) or with an insufficient range of possible duration cue enhancements (see Gordon-Salant, 1986).

A significant group x level interaction [$F(2,27) = 6.64$, $p < 0.01$] indicates that the relative performance of the three groups varied with presentation level. Normal-threshold subjects outperformed both hearing-impaired groups at 75 dB SPL ($p < 0.01$), but both normal-threshold subjects and subjects with gradually sloping losses outperformed subjects with sharply sloping losses at 90 dB SPL ($p < 0.01$). Hearing-impaired listeners probably performed more poorly than

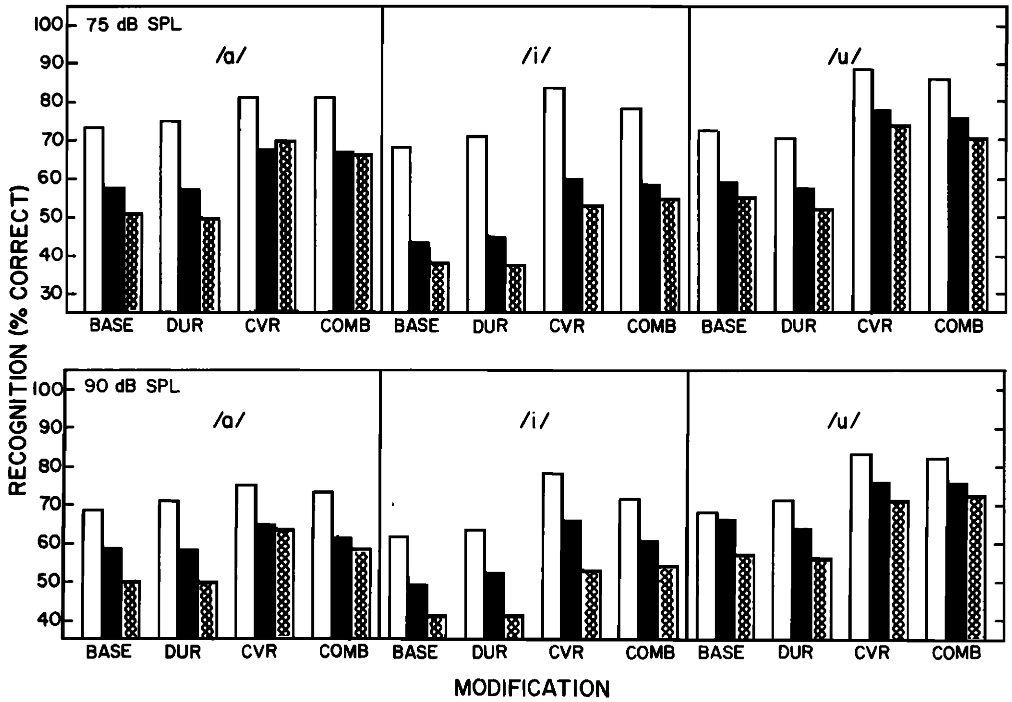

FIG. 1. Percent-correct recognition of nonsense syllables by three subject groups (open bars = subjects with normal hearing, black bars = subjects with gradually sloping losses, transverse bars = subjects with sharply sloping losses), in four acoustic modifications (baseline, duration increase, CVR increase, and combined increase), in three vowel contexts (/a,i,u/) and two presentation levels (75 and 90 dB SPL). The normal-hearing data are taken from Gordon-Salant, 1986.

normal-threshold subjects at the lower signal level because of reduced signal audibility. Consequently, the effective S/B received by the hearing-impaired listeners was reduced, creating a more excessive noise-masking effect. The presence of hearing impairment therefore affected the overall level of performance, although the specific trends in preformance with respect to the acoustic modifications are comparable to those observed from normal-hearing listeners (Gordon-Salant, 1986).

Presentation level also had a different effect on the consonant recognition scores of each group, as shown in Table I. Listeners wtih normal thresholds obtained higher scores at the lower level than at the higher level ($p < 0.01$), listeners with gradually sloping losses obtained higher scores at the higher level than at the lower level ($p < 0.01$), and listeners with sharply sloping losses obtained comparable scores at both levels. The reduced scores of normal-threshold subjects with increasing speech level may be attributed to an overloading of the auditory system. The higher scores of subjects with gradually sloping losses at 90 dB SPL than at 75 dB SPL indicate that signal audibility improved at the higher speech level for these subjects. Increasing speech level did not change scores for listeners with sharply sloping losses, perhaps because the signal audibility did not improve sufficiently in the regions of greatest hearing loss or because of an upward spread of masking effect (Danaher *et al.*, 1973).

TABLE I. Mean percent correct scores, averaged across modification and vowel, of the three subject groups at the two presentation levels.

Group	Level	
	75 dB SPL	90 dB SPL
Normal threshold	77.67	72.33
Gradual slope	60.50	64.25
Sharp slope	56.05	55.75

Specific consonant recognition and confusion patterns were analyzed following procedures detailed in Gordon-Salant (1986). Recognition patterns of the hearing-impaired listeners based on place, manner, and voicing features were comparable to recognition patterns of the normal-hearing listeners (Gordon-Salant, 1986). In addition, the two hearing-impaired groups exhibited similar changes in consonant confusions with acoustic modification as did the normal-hearing subjects (Gordon-Salant, 1986). One striking aspect of these findings was that sizable increases in consonant confusions were associated with the combined increase modification, whereas only minimal consonant confusions were associated with the CVR increase modification.[1]

To summarize, the relative merits of the three kinds of acoustic modifications for normal-hearing listeners reported previously (Gordon-Salant, 1986) were found also in the present study for elderly hearing-impaired subjects. In particular, the CVR increment and combined CVR + duration increment enhanced nonsense syllable recognition and reduced the frequency of major consonant confusions by elderly hearing-impaired subjects. Further, the CVR increment did not produce an increase in other consonant confusions, whereas the combined increment did increase the frequency of other confusions. These findings suggest that little benefit is derived by combining the CVR enhancement with the duration enhancement. The overall improvement in nonsense syllable scores effected by the CVR enhancement was 14%, for both subject groups. These findings are particularly important because increments in performance with the CVR modification were stable at both presentation levels. The principle difference between the present and previous results is seen in the level effects for each of the three groups, which appear to be attributed to the limitations in signal audibility imposed by the listeners' audiometric configurations. Finally, it should be noted that benefits of the CVR increment reported are specific to CV recognition by elderly subjects with the auditory characteristics employed in this study, as well as by young normal-hearing listeners (Gordon-Salant, 1986). Additional research is needed to determine whether such improvements are created by enhancing the CVR in other speech materials, and whether listeners with other audiometric configurations and degrees of hearing loss would exhibit comparable benefits.

ACKNOWLEDGMENTS

This research was supported by a grant from the National Institute on Aging (AG-0065) and by a grant from the General Research Board of the University of Maryland. The computer time for this project was supported in part through the facilities of the Computer Science Center of the University of Maryland. The author is grateful to Tim Bunnell for developing the software used in the experiment, and to Sylvia Campbell and Stacie Raymer for their assistance in data collection and analysis.

[1]For additional details of results from the hearing-impaired listeners, the reader is encouraged to contact the author.

ANSI (**1969**). ANSI S3.6-1969, "American National Standards for Audiometers" (American National Standards Institute, New York).

Chen, F. (**1980**). "Acoustic characteristics and intelligibility of clear and conversational speech at the segmental level," M. A. thesis, MIT.

Danaher, E. M., Osberger, M. J., and Pickett, J. M. (**1973**). "Discrimination of formant frequency transitions in synthetic vowels," J. Speech Hear. Res. **16**, 439–451.

Dubno, J. R., Dirks, D. D., and Morgan, D. E. (**1984**). "Effects of age and mild hearing loss on speech recognition in noise," J. Acoust. Soc. Am. **76**, 87–96.

Dorman, M. F., Marton, K., Hannley, M. T., and Lindholm, J. M. (**1985**). "Phonetic identification by elderly normal and hearing-impaired listeners," J. Acoust. Soc. Am. **77**, 664–670.

French, N. R., and Steinberg, J. C. (**1947**). "Factors governing the intelligibility of speech sounds," J. Acoust. Soc. Am. **19**, 90–114.

Gordon-Salant, S. (**1986**). Recognition of natural and time-intensity altered CVs by young and elderly subjects with normal hearing," J. Acoust. Soc. Am. **80**, 1599–1607.

Jokinen, K. (**1973**). "Presbyacusis VI. Masking of speech," Acta Otolaryngol. **76**, 426–430.

Picheny, M. A., and Durlach, N. I. (**1979**). "Speaking clearly for the hard of hearing," J. Acoust. Soc. Am. Suppl. 1 **65**, S135.