Perceptual and Motor Skills, 1989, 69, 1163-1170.

AGING AND THE INFLUENCE OF CONTEXTUAL CONTRAST ON VOWEL IDENTIFICATION [1]

ROBERT ALLEN FOX AND LESLIE McCRACKEN

The Ohio State University

Summary.—This study examines possible differences between young and older adult listeners in the effect of contextual contrast on vowel identification using a vowel anchoring procedure. All listeners identified a 7-step [i]—[I] continuum under both an equiprobable control condition and an anchoring condition (in which one end-point stimulus occurred four times more often than any other single stimulus token). Phoneme boundaries in the anchoring condition shifted toward the anchor endpoint, as expected, for both groups, but there was a significant effect of age when [i] served as the anchor. Evidence of an increase in response bias for older adult listeners was also found.

A listener's ability to identify an isolated vowel can be affected by the adjacent stimuli in a test sequence. This contextual effect is usually one of contrast and has been demonstrated through many different experimental procedures including paired contrast (Crowder & Repp, 1984), selective adaptation (Morse, Kass, & Turkienicz, 1976) and anchoring (Sawusch & Nusbaum, 1979; Fox, 1985a, 1985b). One can also consider the verbal transformation effect (Warren, 1961), in which a listener's perception of a stimulus token undergoes changes when it is repeated many times in succession, to be another situation in which the context (here, repetitions of the stimulus token itself) significantly affects the perceptual process.

A number of different possible explanations for these contextual effects have been proposed including *feature detector fatigue, response bias* and *changes in auditory ground* (e.g., Fry, Abramson, Eimas, & Liberman, 1962; Ladefoged & Broadbent, 1957; Thompson & Hollien, 1970; Diehl, Elman, & McCusker, 1978; Repp, Healy, & Crowder, 1979; Sawusch & Nusbaum, 1979; Sawusch, Nusbaum, & Swab, 1980; Crowder & Repp, 1984; Fox, 1985a, 1985b). Warren (1985) suggested that all such effects could be subsumed under a more general *criterion shift rule*, which states that the criteria used in evaluating a stimulus may be shifted in the direction of recently or simultaneously experienced stimulus values.

For the most part, the investigations of these contextual effects have

[1]This research was partially supported by Grant 1 R0 AG08353-01 from the National Institute on Aging, National Institutes of Health. An earlier version of this study was presented to the 1988 Annual Convention of the American Speech-Language-Hearing Association, Boston, MA, 18-21 November, 1988. Reprints may be obtained from the first author at the Division of Speech and Hearing Science, The Ohio State University, 324 Derby Hall, 154 N Oval Mall, Columbus, OH 43210-1372.

been limited to young adult listeners and have not addressed changes which the aging process might produce. For example, is there a change—either a decrease or increase—in the extent to which the surrounding auditory/phonetic context influences phonetic identifications as a listener ages? Such changes may provide information about memory changes and/or perceptual processing changes that are partially responsible for the speech comprehension difficulties encountered by the elderly.

There is at least one set of experiments which seems to have shown such differences. While investigating the verbal transformation effect, Warren (1961, 1962) obtained data which demonstrated that young adult listeners were *more* prone to verbal transformations than were elderly listeners. That is, younger listeners heard more illusory perceptual changes in a constantly repeated stimulus than did older listeners. These results could be interpreted as showing that such contextual effects are *reduced* in older adults.

The present study was designed to investigate whether such differences in contextual effects as a function of aging could be obtained using an anchoring procedure. In anchoring experiments listeners identify stimulus tokens under two different conditions. In the baseline condition (equiprobable control condition), the number of presentations of each stimulus token is equal. In the anchoring condition, the anchor stimulus—which is often the endpoint stimulus—is presented many more times than the other stimuli (cf. Sawusch & Nusbaum, 1979; Fox, 1985a, 1985b). The resulting identification scores usually show that identification of the test tokens, particularly the ambiguous ones, are perceived *in contrast to* the more often occurring stimulus. The present experiment will determine whether there are differential vowel identification contrast shifts associated with aging.

Method

The listeners were 30 18- to 25-yr.-old undergraduate students from The Ohio State University and 30 55- to 70-yr.-old individuals who were native American English speakers with no known hearing impairments or neurological abnormalities. All were naive concerning the nature of the experiment. Hearing acuity was screened binaurally at 1000 Hz, 2000 Hz and 4000 Hz at 20 dB for the younger age group and 20 dB, 30 dB and 35 dB, respectively, for the older age group (these different standards reflect expectations of presbycusis in the older age group). These screening procedures ensure that all participants do not have hearing impairments in the range of frequencies most important for the synthetic vowel continua (<4 kHz).

The stimulus set consisted of a seven-step [i]—[I] vowel continuum. The stimuli were constructed using the Klatt (1980) software cascade/parallel speech synthesis program implemented on a PDP 11/23 computer. These steady-state vowels ranged perceptually from [i] as in *beet* to [I] as in *bit* and were modeled after Pisoni's (1971) continuum. Stimulus tokens along the vowel continuum were produced by varying the frequencies of the first three formants as shown in Table 1. The frequencies of Formants 4 and 5 were 3850 Hz and 4900 Hz, respectively, for all stimuli. Each of the stimulus tokens were 420 msec. in duration. All vowels were steady-state (i.e., no change in formant frequencies over time). The fundamental frequency for each token began at 125 Hz and fell linearly, reaching 100 Hz after 360 msec. where it remained at 100 Hz until the end of the token. All vowels were digitized at a 12-bit quantization level, low-pass filtered at 4.8 kHz and output at a 10-kHz sampling rate.

TABLE 1
FORMANT FREQUENCY VALUES USED IN SYNTHESIS OF VOWEL CONTINUUM
(ALL VALUES IN HZ)

Stimulus Step	F1	F2	F3
1	269	2296	3019
2	285	2263	2955
3	297	2230	2912
4	315	2183	2829
5	336	2151	2769
6	354	2105	2709
7	375	2075	2670

These stimuli were converted to analog form to make a baseline tape and two different anchor tapes. The baseline tape served to determine the identification function of the vowels in the continua when all tokens were equally likely to occur. In the baseline condition the stimulus tape contained 20 repetitions of each of the seven test tokens in random order, for a total of 140 tokens. Each of the anchor tapes had 80 occurrences of one of the endpoint stimuli (either step 1—[i] or step 7—[I] of the continuum) together with 20 occurrences of the remaining six vowel stimuli for a total of 200 tokens. On each of the anchor tapes the order of the stimuli was randomized, with the restriction that no single stimulus could occur more than three times in succession. All three tapes were recorded with 4-sec. interstimulus intervals.

Each listener was tested individually in one 45-min. session. The stimulus tapes were reproduced on a high-quality stereo cassette recorder (Sony TC-FX 705) and were presented to the listeners at a comfortable listening level (approximately 75 dB SPL) via headphones (Sennheiser HD 420). The listeners in each age group were randomly assigned to one of the two anchor conditions. Thus each listener listened to the baseline-tape first and then to either the [i]- or [I]-anchor tape. The listeners were told that they would be listening to speech tokens that would sound like the vowel in *heed* or the vowel in *hid*. They were warned that the stimuli were isolated vowels and that they would not hear the "h" or the "d." Listeners were asked to identify the test tokens as either [i] (*heed*) or [I] (*hid*) by circling the appropriate response on their answer sheets.

Listeners were also required to rate the confidence that they had in their identification choice. A 4-point scale was used in the confidence rating, with 1 indicating that the listener was positive his identification response was correct, 2 indicating his response was probably correct, 3 indicating his response was possibly correct, and 4 indicating that his response was a guess. A 3-min. break occurred between the control tape and the anchor tape. There were no new instructions given at this time concerning any differences between the two tapes.

Results

All identification and rating responses were converted into an 8-point scale (Sawusch & Nusbaum, 1979), with 1 indicating a very positive "heed" response and 8 representing a very positive "hid" response. The "phoneme boundaries" (the hypothetical point in the continuum which would have a mean of 4.5) for each listener was determined using linear interpolation between points on both sides of the boundary in both the control and anchor conditions. The obtained phoneme boundaries and boundary shifts are shown in Table 2.

TABLE 2
Mean Phoneme Boundaries and Boundary Shifts For Both Age Groups

	Baseline	Anchor	Shift (baseline-anchor)
[i]—anchor condition			
Young	4.07	3.57	0.50†
Old	4.63	3.58	1.04‡
[I]—anchor condition			
Young	4.24	4.60	−0.36‡
Old	4.33	4.61	−0.27*

$^*p<.05$. $^\dagger p<.01$. $^\ddagger p<.001$.

For both age groups the /i/-anchor produced a significant shift in the phoneme boundary in the direction of the anchor vowel at the .001 level (two-tailed t test). This shift indicates that the vowels are being identified in *contrast* to the more often occurring anchor vowel. That is, the vowel stimuli are being perceived more often as /I/ in the anchor condition than in the baseline condition. The mean boundary shift for the young adult listeners was 0.50 ($t_{14} = 3.94$, $p<.001$) and 1.04 ($t_{14} = 6.08$, $p<.001$) for the older adult listeners.

Significant boundary shifts in the opposite direction were obtained using the /I/ as anchor. The mean boundary shift for the younger listeners was -0.36 ($t_{14} = 4.56$, $p < .001$). The mean boundary shift for the older listeners was slightly less at -0.27 ($t_{14} = 2.27$, $p < .04$). In both age groups, the boundary shifts represented significant contrast effects. The boundary shifts were somewhat greater for the /i/ anchor than for the /I/ anchor, which is consistent with the findings of Sawusch, *et al.* (1980).

In terms of age-related differences, it would seem that a greater boundary shift occurred in the older age group with the [i]-anchor. The boundary shift data were analyzed using a one-way analysis of variance using the factor age (young adult and older adult) in each of the two anchor conditions. These data represent the boundary shift obtained from each separate listener (which was the baseline boundary minus the anchor boundary). In the [I]-anchor condition, there was no significant main effect for age ($F_{1,29} = 0.35$, $p < .56$). However, for the [i]-anchor condition, there was a significant main effect for age ($F_{1,29} = 6.55$, $p < .02$).

Discussion

The basic results of the present study agree with the findings of Sawusch and Nusbaum (1979) and Fox (1985a, 1985b) in that contrast effects were found which produced significant category boundary shifts in the anchor condition relative to the equiprobable control. For both age groups there was a difference in the magnitude of the shift for the [i] anchor as opposed to the magnitude of the shift for the [I] anchor in the older age group.

In terms of age-related effects, however, the results were opposite to those that might have been expected, given Warren's (1961, 1962) verbal transformation data. In particular, Warren (1961) found that young adults were more likely to experience verbal transformations than older adults. If one views the verbal transformation effect in terms of contextual contrast and/or assimilation effects, then one might assume that older listeners would be *less* sensitive to the auditory/phonetic context in making an identification. Were this the case, we would expect the boundary shifts of the older listeners to be smaller than those for the younger listeners, a result opposite to the one obtained.

However, the results of the present study demonstrate that the identifications of older listeners were affected by the context *as much as* or *more than* were the younger listeners. There is no evidence of a decline in the representation of speech signals in auditory memory or its role in vowel identification, which is consistent with many studies (e.g., Craik, 1977; Arenberg, 1976; Parkinson & Perey, 1980; but note Crowder, 1980). How, then, does one account for the differences between these data and those of Warren (1961)?

At least part of Warren's results can be explained by response differences or bias in young vs older adult subjects. In particular, several studies (including Wallach & Kogan, 1961; Botwinick, 1969; Marshall, 1981) have demonstrated that elderly subjects behave more cautiously than do young adult subjects in an experimental situation although Gordon-Salant (1986) has shown this to be task-dependent to some extent. This conservatism is especially evident when subjects are allowed the avoidance option, that is, they are allowed not to respond, and when the response set is open rather than closed (Gordon-Salant, 1986). In Warren's (1961) study, the "avoidance" response would be a failure to indicate that a perceptual change had occurred in the continuous repetition of a known stimulus. The pattern of Warren's data does not necessarily indicate that elderly listeners are less affected by contextual influences; it simply does not indicate a perceptual change has occurred. It is likely that the verbal transformation task does not allow the full extent of contextual effects to be ascertained.

To explain the significantly greater boundary shifts found in the older listeners in the [i]-anchor condition, one is tempted to pursue Sawusch, *et al.*'s (1980) speculation about the nature of anchoring with the [i] vowel and discuss possible age-related differences in the access of phonetic prototypes from long-term memory.[2] However, a more likely explanation would also involve a type of response bias.

In particular, listeners in the experimental task were hearing isolated vowels from a vowel continuum which changed only in terms of phonetic vowel quality (from [i] to [I]). However, there is a phonological difference between [i] and [I] which might produce perceptual differences on identification tests. In English, lax vowels (such as [I]) are not allowed in open syllables whereas tense vowels (such as [i]) are allowed (Ladefoged, 1982). Thus the [i] vowel may be able to establish a more salient auditory/phonetic ground against which ambiguous tokens may be contrasted than does the [I] vowel because the [i] is phonologically acceptable in the form in which it was presented. If older adult listeners are more conservative and rely on this phonological constraint more than do younger adult listeners, they might be more prone to such an effect. This could explain why in the [i]-anchor condition the boundary shift for the older listeners was significantly greater than those for the younger listeners, while in the [I]-anchor condition the shift was smaller (though not significantly).

The present study provides evidence that older listeners are not significantly different from younger listeners in terms of the influence of phonetic context upon vowel identification but differences in response biasing may be

[2]Sawusch, *et al.* (1980) speculated that anchoring with the [I] vowel was mediated by auditory memory, whereas anchoring with the [i] vowel, which is a "point" vowel, was related to early perceptual processing which might involve the retuning of a vowel prototype from long-term memory.

present. Additional research must address whether these effects can be extended to other vowel continua as well as to consonant stimuli.

REFERENCES

ARENBERG, D. (1976) The effects of input condition on free recall in young and old adults. *Journal of Gerontology*, 31, 551-555.

BOTWINICK, J. (1969) Disinclination to venture response versus cautiousness in responding: age differences. *Journal of Genetic Psychology*, 113, 55-62.

CRAIK, F. I. M. (1977) Age differences in human memory. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging*. New York: Van Nostrand Reinhold. Pp. 384-420.

CROWDER, R. G. (1980) Echoic memory and the study of aging memory systems. In L. W. Poon, J. L Fozard, L S Cermak, D. Arenberg, & L. W. Thompson (Eds.), *New directions in memory and aging* Hillsdale, NJ: Erlbaum. Pp. 181-204.

CROWDER, R. G., & REPP, B H. (1984) Single formant contrast in vowel identification. *Perception & Psychophysics*, 35, 372-378.

DIEHL, R. L., ELMAN, J. L., & MCCUSKER, S. B. (1978) Contrast effects in stop consonant identification. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 599-609.

FOX, R. A. (1985a) Auditory contrast and speaker quality variation in vowel perception. *Journal of the Acoustical Society of America*, 77, 1552-1559.

FOX, R. A. (1985b) Within- and between-series contrast in vowel identification: full-vowel versus single-formant anchors. *Perception & Psychophysics*, 38, 223-226.

FRY, D. B., ABRAMSON, A. S., EIMAS, P. D., & LIBERMAN, A. M. (1962) The identification and discrimination of synthetic vowels. *Language and Speech*, 5, 171-189.

GORDON-SALANT, S. (1986) Effects of age on response criteria in speech-recognition tasks. *Journal of Speech and Hearing Research*, 29, 155-162.

KLATT, D. H. (1980) Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America*, 67, 971-995.

LADEFOGED, P. (1982) *A course in phonetics*. New York: Harcourt Brace Jovanovich.

LADEFOGED, P., & BROADBENT, D. E. (1957) Information conveyed by vowels. *Journal of the Acoustical Society of America*, 29, 98-104.

MARSHALL, L. (1981) Auditory processing in aging listeners. *Journal of Speech and Hearing Disorders*, 46, 226-240.

MORSE, P. A., KASS, J. E., & TURKIENICZ, R. (1976) Selection adaptation of vowels. *Perception & Psychophysics*, 19, 137-143.

PARKINSON, S. R., & PEREY, A. (1980) Aging, digit span and stimulus suffix effect. *Journal of Gerontology*, 35, 736-742.

PISONI, D. B. (1971) On the nature of the categorical perception of speech sounds. Unpublished doctoral dissertation, Univer. of Michigan.

REPP, H. B., HEALY, A. F., & CROWDER, R. G. (1979) Categories and context in the perception of isolated steady-state vowels. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 129-145.

SAWUSCH, J. R., & NUSBAUM, H. C. (1979) Contextual effects in vowel perception: I. Anchor-induced contrast effects. *Perception & Psychophysics*, 25, 292-302.

SAWUSCH, J. R., NUSBAUM, H. C., & SCHWAB, E. C. (1980) Contextual effects in vowel perception: II. Evidence for two processing mechanisms. *Perception & Psychophysics*, 27, 421-434.

THOMPSON, C. L., & HOLLIEN, H. (1970) Some contextual effects on the perception of synthetic vowels. *Language and Speech*, 13, 1-13.

WALLACH, M. A., & KOGAN, N. (1961) Aspects of judgments and decision making: interrelationships and changes with age. *Behavioral Science*, 6, 23-36.

WARREN, R. M. (1961) Illusory changes in repeated words: differences between young adults and the aged. *American Journal of Psychology*, 74, 506-515.

WARREN, R. M. (1962) An example of more accurate auditory perception in the aged. In C.

Tibbitts & W. Donahue (Eds.), *Social and psychological aspects of aging.* New York: Columbia Univer. Pp. 789-794.

WARREN, R. M. (1985) Criterion shift rule and perceptual homeostasis. *Psychological Review*, 92, 574-584.

WORKING GROUP ON SPEECH UNDERSTANDING AND AGING (National Research Council). (1987) Speech understanding and aging. *Journal of the Acoustical Society of America*, 83, 859-895.

Accepted December 4, 1989.