Dissociations between Vowel Durations and Formant Frequency Characteristics

Article in Journal of speech and hearing research · July 1985
DOI: 10.1044/jshr.2802.255 · Source: PubMed

Nan Ratner
University of Maryland, College Park

Journal of Speech and Hearing Research, Volume 28, 255–264, June 1985

DISSOCIATIONS BETWEEN VOWEL DURATIONS AND FORMANT FREQUENCY CHARACTERISTICS

NAN BERNSTEIN-RATNER
University of Maryland, College Park

Disagreement exists on the degree to which rate of speech and segmental duration affect the formant frequency characteristics of vowels. Post hoc analysis of the vowel characteristics of words uttered by women in conversational speech with both adult and child addressees indicates that there is no simple relationship between the length of vowels and the degree to which their formant frequency characteristics resemble those seen in citation forms of speech. In the case of women addressing children, it was possible for content and function words to share formant frequency characteristics that maximally differentiated their embedded vowels, despite the relatively shorter duration of function word vowels. Implications for the elicitation of "clear speech" are discussed.

Disagreement exists regarding the effects of rate of speech and of segmental duration on vowel formant frequency values (Miller, 1981). Tiffany (1959) noted that, as speakers read from text at a conversational speaking rate, reduced contrast among vowel formant frequency characteristics distinguished their productions from those observed when they read single words in isolation. A few years later, Lindblom (1963) described a tendency for text-read speech uttered at various rates to be characterized by undershoot of vowel formant frequency values (centralization) as rate increased, and concluded that vowel "duration seems to be the main determinant of the reduction" (p. 1780). Delattre (1969) performed cross-linguistic analyses of vowel reduction and suggested that data from languages other than English indicate that stress, rather than duration, might be the factor primarily responsible for changes in vowel formant frequency characteristics in varying contexts; because the two are closely correlated functions in spoken English, the attribution of vowel centralization simply to decreases in vowel duration might be spurious. More recently, Gay, Ushijima, Hirose, and Cooper (1974) noted significant vowel undershoot at experimentally induced rapid rates of speech. Koopmans-van Beinum (1980) noted high correlations both between rate and vowel duration and between vowel duration and vowel centralization. However, high correlations were not uniformly observed for all vowels analyzed, and all correlations were lower when conversational speech, rather than text, was analyzed.

Conversely, a number of studies have suggested that vowel duration and formant frequency characteristics do not appear to be systematically related. Gay (1978) found that, when speakers were instructed to change speech rate but maintain phoneme identities ("be clear"), vowel formant frequencies remained relatively stabile, though vowel durations were compressed. Verbrugge and Shankweiler (1977) noted a similar finding: Although vowel durations decreased during fast speech conditions, formant frequencies were only slightly centralized.

Accompanying disagreement regarding the effects of rate and duration on the acoustic characteristics of vowels have been equivocal findings on vowel intelligibility under conditions of rate and durational manipulation. In a study somewhat like Gay's (1978) investigation, Tollhurst (1957) found that instructions to speakers to either decrease rate or increase intelligibility resulted in increased intelligibility. House, Williams, Hecker, and Kryter (1965) noted, however, that as one of their speakers lengthened his vowel productions during the production of a speech discrimination stimulus word list, listeners' performance on that speaker's word list fell. Picheny (1981) found that, whereas formant frequency characteristics of vowels contributed to intelligibility ratings, durational characteristics apparently did not, and that no systematic relationship could be observed between vowel durations and vowel formant frequency values.

The present report indicates that the degree to which a vowel's formant frequencies approximate those seen in citation forms of speech, or the degree to which they are centralized, may be partly independent of vowel duration, at least for the specific conversational register explored. Although originally designed to describe the acoustic-phonetic characteristics of the mother–child (or input) speech register (Bernstein-Ratner, 1984a, 1984b), the study provided ancillary evidence of an apparent dissociation between segmental duration and vowel formant characteristics. Implications of such a possible dissociation for the elicitation of "clear speech" register and stimuli are discussed.

METHOD

Subjects

Subjects were five mothers of female children enrolled in a longitudinal and cross-sectional investigation of the acoustic-phonetic characteristics of the mother–child speech register (Bernstein-Ratner, 1984a, 1984b). All mothers were speakers of Standard American English

0022-4685/85/2802-0255$01.00/0

dialects. For purposes of the original studies, children had been grouped in various groups based on their expressive language abilities. This study reports on the behaviors of the mothers of the children in only one of these groups. Each mother in this particular group had a daughter who demonstrated the ability to use simple combinatorial language (MLUs of approximately 2.0–3.0) and ranged in age from 1:5 to 1:8 (years:months).

Data Collection

The conversations of mother–child dyads were individually recorded in a sound-proofed playroom. Mothers were told only that the study was concerned with the linguistic interaction of mothers and language-learning children. Mothers wore Sony ECM-50 lavalier microphones that had long, flexible leads, and they were informed that they were serving as "microphone holders" for their children because safety concerns made placement of microphones on their infants undesirable. The women were simply instructed to play naturally with the children using an assortment of toys, and to maintain a fairly small distance between themselves and their daughter to ensure recording of the child's language.

Play sessions, which lasted approximately 30 min, were carried out three times over a 6-month period. Following each session, mothers were individually interviewed by the investigator. Although ostensibly designed to elicit the mothers' opinions about their children's play and language, such interviews actually served to obtain a sample of each woman's speech to an adult addressee.

Following tape transcription, the mother–child and mother–investigator conversations were analyzed to locate instances of monosyllabic words used by each woman to both her child and the investigator. Certain contextual constraints were placed upon the selection of such matched words for analysis. These constraints included the exclusion of isolated single words addressed to the child. Words embedded in questions were never matched with words embedded in declarative intonation contours. Words taken from utterance-final position in speech to one addressee were always matched with the identical word in the same sentence position to the opposite addressee, as utterance-final segments are characterized by increased vowel duration (Klatt, 1975). An array of content and function words representing a range of monophthong and diphthong English vowels was selected. Following procedures described in the next section, the final body of words analyzed for this report consisted of 430 words addressed to the children and 520 words addressed to the adult listener. Of the body of words addressed to the children, 61% were content words (nouns, main verbs, adverbs, and adjectives) and 39% were function words (pronouns, deictics, auxiliaries, modals, prepositions, and clause headers). Of the words addressed to the adult, 60% were content words and 40% were function words.

Analyses

A waveform editing program was used to excerpt target words from the audiotaped interactions for further acoustic analysis. Excerpted words were then subjected to both formant frequency analysis and to measurement of the duration of their embedded vowels. Analyses were performed with the aid of two computer programs. One program performed waveform display and permitted vowel isolation, truncation, and measurement to the nearest millisecond. The second program utilized linear prediction to display the first five formants of each token word in a computerized spectrographic analysis.[1] Frequency values for formants were selected and printed by the program. Utilizing stability, centrality, and peak amplitude cues, a measurement point was selected within the vowel nucleus, and the computer's values for that point in time recorded. Stability was defined as lack of formant "slope" either upward or downward for a period of three sampling points (30 ms) and as relative continuity of the computer plot for an equivalent period of time (i.e., no scattering or absence of plotted values.) Centrality was a simple estimate of the temporal midpoint of the truncated vowel. Peak amplitude of the syllable was defined as that point in time when the intensity of the vowel segment reached its maximum value. Quite often, these three criteria could be satisfied simultaneously. If they could not, a two-out-of-three criterion was employed. Occasionally, the program would plot multiple formant tracks where only one would have been predicted given the target vowel. Thus, all values printed by the program were checked against known norms for women's vowel production (Peterson & Barney, 1954) to select the most reasonable value for the target vowel in such cases.

A small subset of the data (40 tokens) was reanalyzed separately by the investigator to appraise reliability of the measurement criteria. Mean disagreement in Hz for the obtained F_1 and F_2 values was ascertained. Mean absolute measurement error for F_1 and F_2 was computed at 38.3 Hz and 75.7 Hz, respectively. Such levels of error

[1]Sampling rate for both programs was 10 kHz. The predictor order for the formant plotting program was 14. Although this is a lower value than that used by Monsen and Engebretson (1983) in their study of the accuracy of formant frequency measurements made by linear prediction, higher orders tended to produce multiple extraneous formant tracks, especially in the mother–child speech register. Typically, this register is characterized by a relatively high frequency fundamental. In the Monsen and Engebretson study, which analyzed synthesized vowels with known characteristics, an order of 14 was not chosen; for fundamentals approximating those utilized by the women in this study (200–250 Hz; Bernstein-Ratner & Pye, 1984), their utilization of an order of prediction of 16 produced an average F_1 error of −113 Hz and an average F_2 error of only +9 Hz; at an order of analysis of 12, F_1 errors decreased to −34 Hz, but F_2 errors increased to +557 Hz. Since the program utilized in the present investigation operated on an interpolated order of analysis, accuracy of its "extraction" properties cannot be estimated from the previous study. However, since all tokens were analyzed in the same manner, measurement tendencies of the program should have operated equally across addressee groups and affected all measurements equally.

may be compared with those obtained by Monsen and Engebretson (1983). They concluded that linear prediction of speech having fundamental frequency characteristics of approximately 200–300 Hz (comparable to the fundamentals of the women in the present study; Bernstein-Ratner & Pye, 1984) was accurate to approximately 61–67 Hz, when results obtained by linear prediction are compared against the known characteristics of synthesized vowel tokens. Thus, the level of measurement reliability in this study can be taken to be approximately equal to the inherent level of error that accompanies the decision to employ linear prediction as a method of analysis.

Data for each token were entered into computer library files for further statistical and descriptive treatment. In an effort to guard further against measurement or recording error, procedures Broad and Wakita (1977) used were employed to eliminate from analysis all tokens whose F_1 or F_2 values fell beyond two standard deviations from the calculated mean frequency values for each vowel analyzed in each data set. This procedure resulted in the loss of 9.3% of the tokens originally selected for analysis and left the 950 tokens outlined in the previous section.

RESULTS

Figures 1 and 2 plot the formant frequency characteristics of vowels the women uttered to the investigator and to the child addressees. Characteristics of content and function words are plotted separately. Because vowels in function words are typically shorter in duration and relatively destressed when compared to vowels in content words, they might be expected to demonstrate a more centralized formant frequency pattern than content-word vowels. Figure 1, which depicts the formant frequencies of words of adult speakers to an adult listener, supports such a hypothesis quite clearly. Although the "point vowels" /i, a, u/ are relatively unaffected by part of speech, the remainder of the vowels embedded in function-word contexts show a noticeable tendency towards more central first and second formant positions. A series of t tests for differences between independent samples was carried out to compare formant frequency values of the vowels across part of speech. Because such comparisons involved eight vowels on two measures (F_1, F_2), the alpha level for the test of this hypothesis was set at $p \leq .0032$ (.05/16) (Kirk, 1969). Two vowels demonstrated significant part-of-speech effects, both for second formant frequency characteristics: /ɛ/ and /ʊ/[2] [$t(15) = 4.0659$, $p \leq .0010$; $t(38) = -3.2364$, $p \leq .0025$, respectively].

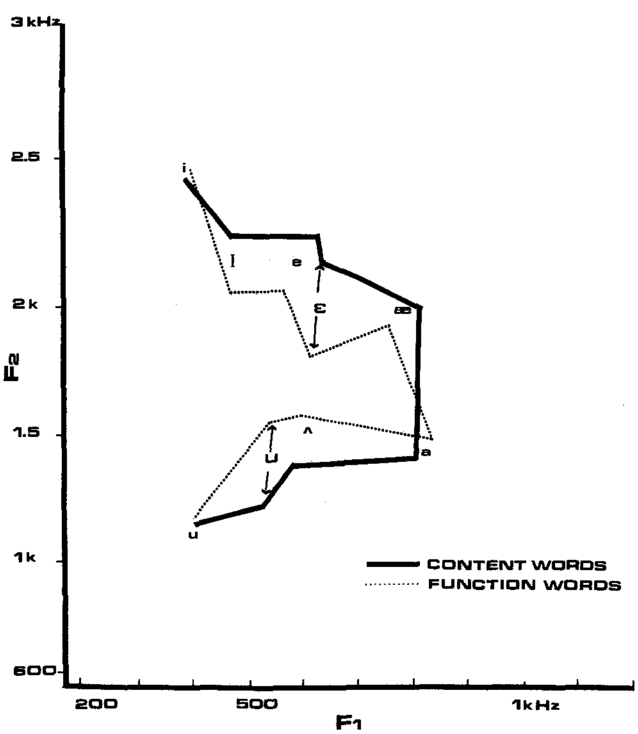

FIGURE 1. First (F_1) and second (F_2) formant frequency characteristics of content- and function-word vowels uttered by women to an adult listener. Arrows indicate differences significant at the $p \leq .003125$ level.

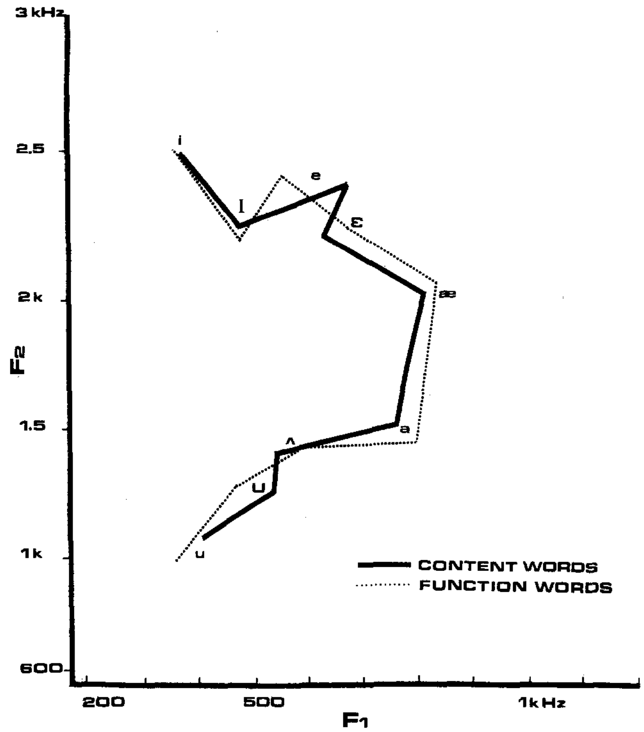

FIGURE 2. Formant frequency characteristics of content- and function-word vowels uttered by women to child listeners.

However, vowels in maternal speech to children seem to be relatively unaffected by part of speech, both subjectively and by statistical test. Figure 2 provides an illustration of the formant frequencies of vowels in words spoken to children. No comparisons reached statistical significance at $p \leq .0032$ for this speaker condition. Thus, the

[2]Formant frequency characteristics and durational properties are provided for each vowel for each addressee condition in the Appendixes. Additionally, t-test comparison results for the variables F_1, F_2 and duration by addressee and by grammatical class are provided.

same group of women, producing matched groups of words to both adult and child addressees, appeared to reduce or centralize function-word vowels when conversing with mature conversational partners. Child addressees were not the recipients of reduced vowels even in small grammatical morphemes. This style shift may be observed more clearly if one overlays content-word articulation and function-word articulation to the two groups of addressees, as has been done in Figures 3 and 4.

Figures 5 and 6 display the durational characteristics of the vowels plotted in Figures 1 and 2. Generally, mothers' speech to an adult addressee (Figure 5) was characterized by shorter vowel duration in function words—not an unexpected finding. In two cases, /æ/ and /u/, a shortening of vowel duration in function-word environments was statistically significant at $p \leq .00625$ (.05/8 comparisons) [$t(56) = 2.8782$, $p \leq .0057$; $t(36) = 5.4115$, $p \leq .0001$, respectively]. For five of the six remaining vowels, shortening in function-word contexts was simply an observable trend. Only in the case of /i/ did function-word context not result in relative vowel shortening.

Speech to a child listener (Figure 6) was characterized by similar durational properties. For six of the eight vowels, function-word contexts were associated with shorter vowel durations than were content-word environments; this trend reached significance with /a/ [$t(32) = 3.1029$, $p \leq .0046$]. As in the case of speech to the investigator, however, there were exceptions. In the child-addressed sample, it was the vowels /u/ and /ʊ/ that were relatively lengthened in function words. /ʊ/ was lengthened by 10 ms and /u/ by 23 ms in function-word environments in mother–child speech. For /u/, such a trend contrasts directly with the statistically significant shortening which characterized this vowel in adult–adult function words. Because the same few types of words were being sampled in the same relative utterance positions, it is unclear why mothers changed the durational properties of this vowel across addressees.

Whereas function words in adult–adult speech were both shorter and centralized, function-word vowels in child-addressed speech were shorter but not centralized. Only in the case of child-directed function-word /u/ can an argument be made for a positive relationship between duration and more extreme formant frequency achievement. This single exception would seem to be directly contradicted by the case of function-word /ʊ/, which was longer but more centralized than its content-word variant in speech addressed to children. To evaluate the possibility that child-addressed speech is characterized by vowels which are globally longer than those in adult-addressed speech, and therefore less likely to be centralized, the durations of adult- and child-addressed content-word vowels and adult- and child-addressed function-word vowels may be superimposed in much the same way that formant frequency characteristics were displayed in Figures 3 and 4.

Figure 7 displays the durations of vowels in content words as spoken to adult and child addressees. There is little observable difference in their durations, and no

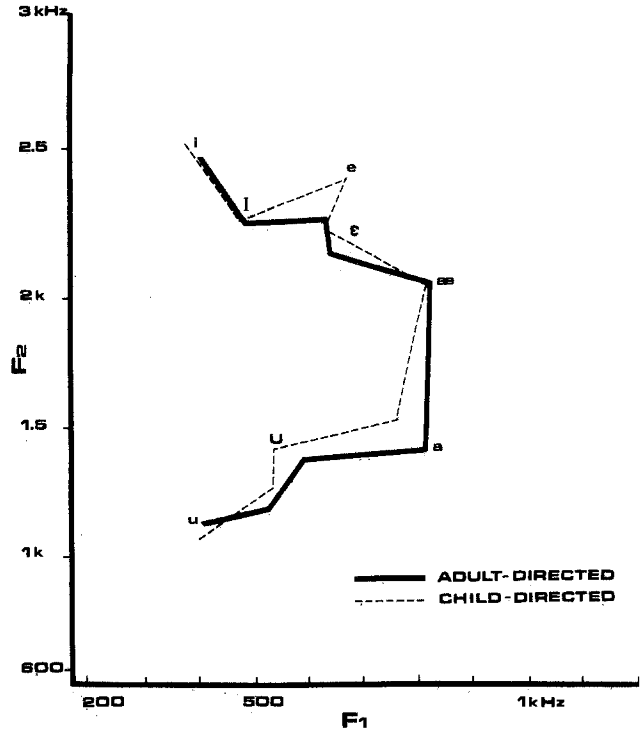

FIGURE 3. Content-word vowels spoken to adult and child listeners.

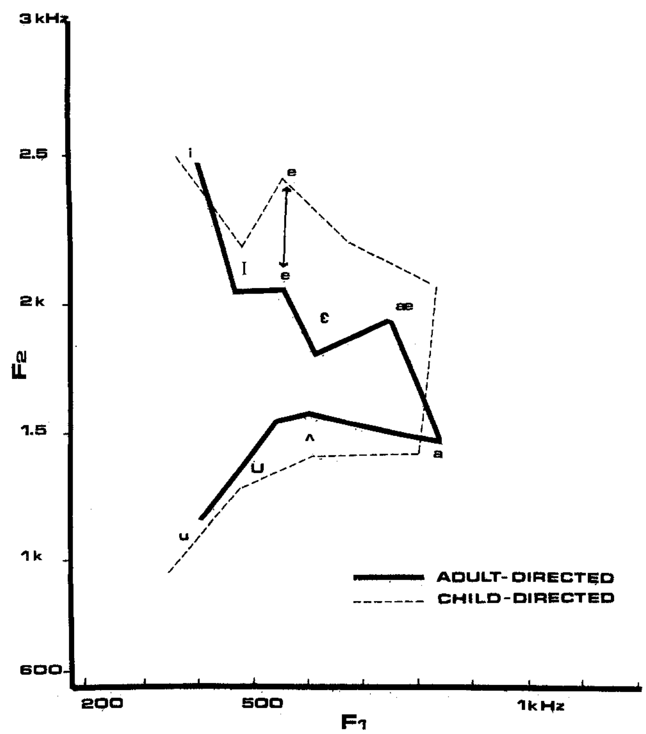

FIGURE 4. Function-word vowels spoken to adult and child listeners. Arrow indicates difference that reached significance at $p \leq .00325$.

within-vowel comparisons across addressee groups are significant at the $p \leq .00625$ level. Thus, mother–child speech was not characterized by generally longer vowel segments, even in lexical environments, although the

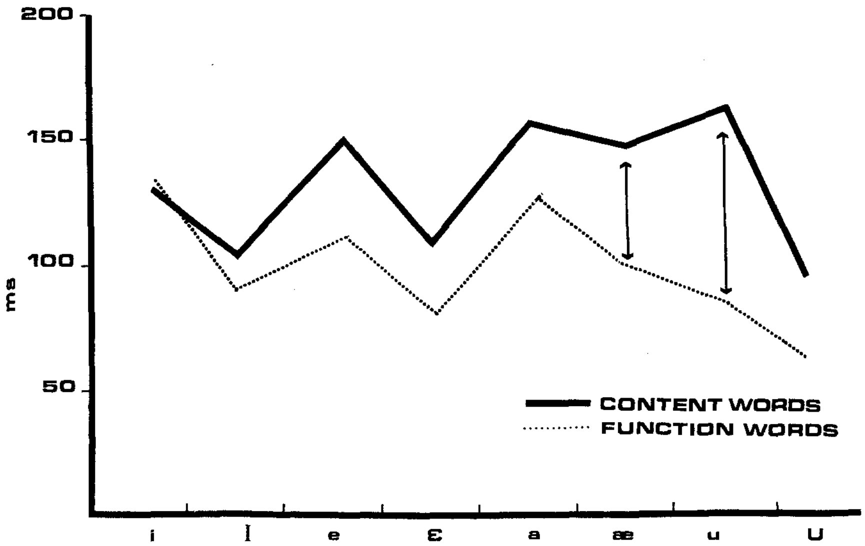

FIGURE 5. Durations of content- and function-word vowels spoken by women to an adult listener. Arrows indicate differences that reached significance at $p \leq .00625$.

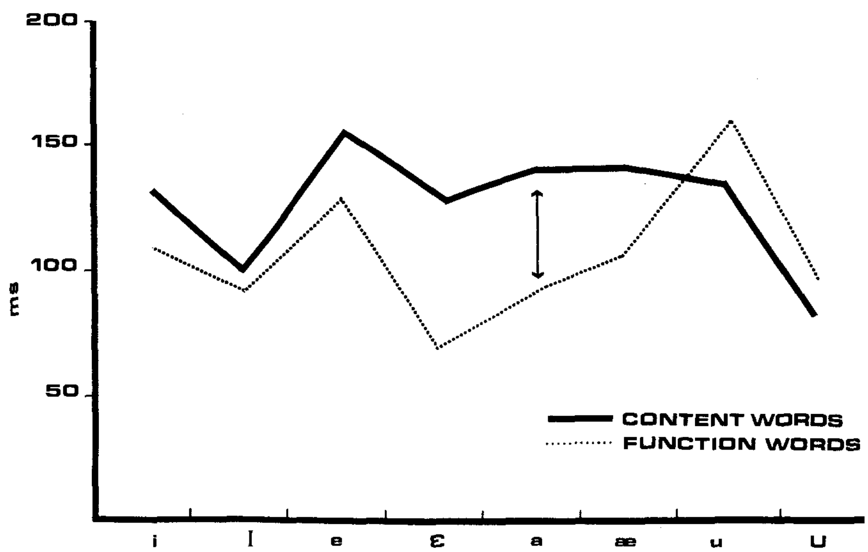

FIGURE 6. Durations of content- and function-word vowels spoken to child listeners. Arrow indicates difference significant at $p \leq .00625$.

shorter utterance length of mother–child speech might have predicted relatively longer segmental durations in this register. Only /i/, /e/, and /ε/ are somewhat longer in child-addressed speech; for /ε/ this difference is marked and does coincide with relatively less centralized production of this vowel.

Figure 8, which displays durational data for function-word vowels to the two addressee groups, shows a more varied pattern. Four of the eight vowels are longer to adult addressees; the other half are longer to child addressees. The sole comparison to reach significance at the $p \leq .00625$ level is /u/, which is significantly longer in function words uttered to children than in function words in adult–adult speech [$t(25) = -3.0554$, $p \leq .0053$].

This significant difference in vowel duration coincides with observable differences between addressee groups for the formant frequency properties of /u/ in function word environments (Figure 4), although the formant frequency difference is not statistically significant at the conservatively set alpha level. For /u/ it may be appropriate to hypothesize that its significantly longer duration in child-addressed speech plays a role in its relatively more extreme formant frequency appearance.

For four of the eight vowels, there is a negative or nonsystematic relationship between relative vowel duration and formant frequency properties. /ɪ/ and /æ/ were of identical duration in function words to child and adult addressees, yet both were more centralized in adult–adult speech than in mother–child speech. /i/ and /ε/ were shorter in mother–child than in adult–adult conversational speech, but were less centralized in the former register. For /a/, rather substantial shortening of child-ad-

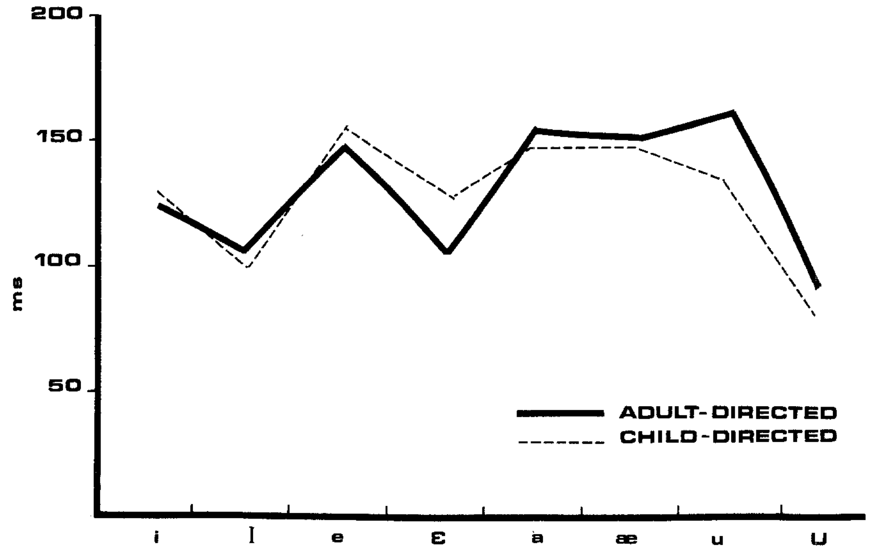

FIGURE 7. Durations of content-word vowels spoken to adult and child listeners.

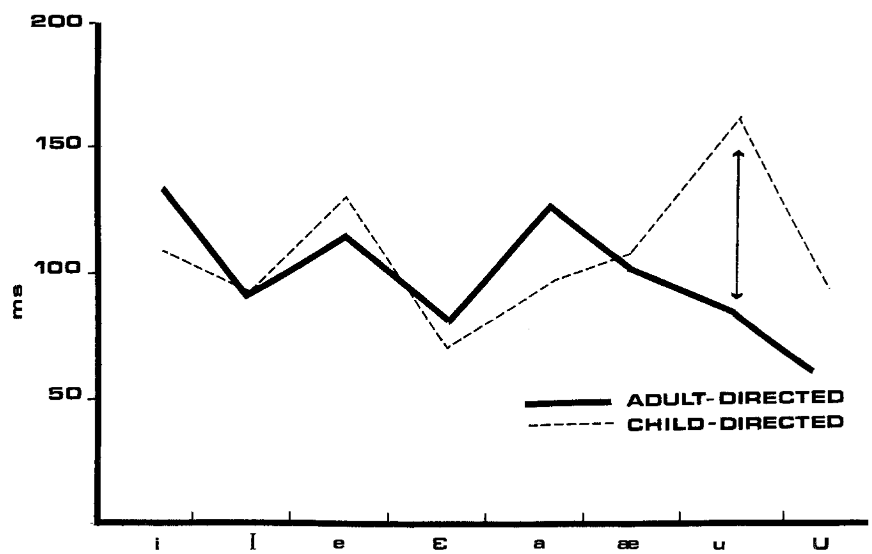

FIGURE 8. Durations of function-word vowels spoken to adult and child listeners. Arrow indicates difference significant at $p \leq .00625$.

dressed vowels did not appreciably change that vowel's formant frequency characteristics in the two registers. The remaining two vowels, /e/ and /ʊ/, were longer in function words spoken to children, and were also less centralized in that register.

Formant frequency and durational properties of diphthong vowels were also plotted for adult–adult and mother–child speech; these data are displayed in Figures 9 and 10. As in the case of monophthong vowels, one fails to see a clear relationship between segmental duration and formant frequency characteristics. Briefly, there is a nonsignificant trend for child-addressed diphthongs to show more extreme excursion between onset and offglide values than adult-addressed diphthongs. However, the four diphthongs analyzed are of almost identical duration for both addressee groups.

DISCUSSION

Although based on post hoc analysis, the findings of this study suggest that no simple relationship exists between the length of vowel segments and their formant frequency characteristics in spontaneous conversational speech. In this, the present findings are in concordance with those obtained by Verbrugge and Shankweiler (1977) and by Gay (1978), whose subjects' vowel productions during fast reading showed little effect of duration on formant frequency characteristics. To a lesser degree, these results are also congruent with those reported by Koopmans-van Beinum (1980). Her analyses of the structured and conversational speech of men and women indicated nonuniform correlations between vowel duration and vowel centralization patterns. The magnitude of

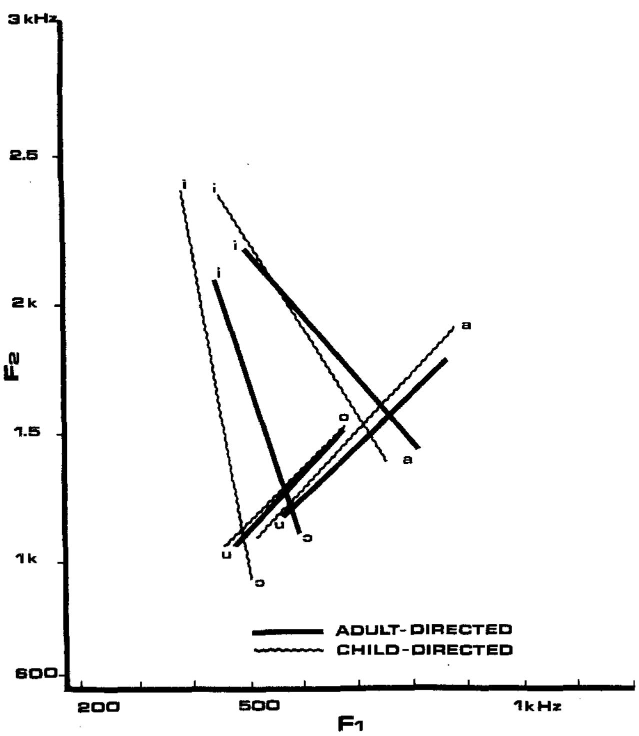

FIGURE 9. Formant frequency characteristics of diphthongs spoken to adult and child listeners.

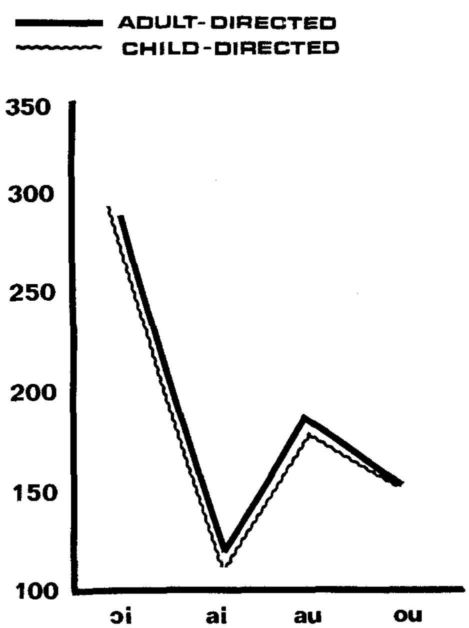

FIGURE 10. Durations of diphthongs spoken to adult and child listeners.

observed correlations fell as her subjects' tasks changed from reading of words and carrier phrases to conversational speech.

There are a number of possible explanations for such a dissociation between vowel duration and formant frequency values, especially in the mother–child register examined in this study. As discussed in greater detail in Bernstein-Ratner (1984a), it is not altogether illogical for women who are addressing those children who are learning language to place some degree of atypical emphasis on the articulation of function words. Although such emphasis would not serve any particular function for speech addressed to linguistically proficient adults, or for children who have not yet begun to use recognizable words, it could serve an important teaching function for children who have a core expressive vocabulary and who have learned to combine words to express basic semantic relationships. Such children are essentially on the brink of developing grammar and will need to develop an increased awareness of those classes of words (such as articles, auxiliaries, pronouns, etc.) that elaborate the message and make it adult-like in syntactic form. Malsheen (1980) reported similar findings for VOT in mother–child speech; that is, maternal function-word articulation to language-learning children was characterized by more precise VOT values than was articulation of function words to adult conversational partners.

Although it is unreasonable to assume that the mothers of such children make a conscious effort to clarify grammatical functors, it is interesting that Gay (1978) and Picheny (1981) found that instructions to "be clear" or to "maintain clarity" resulted in vowel articulation which, although shorter in duration during rapid speech, was quite consistent in formant frequency characteristics. Thus, mothers, who may have a subconscious awareness of the need to provide their children with input conducive to processing and learning speech may access a "clarity" register that maximizes phonemic contrasts across the board.

Picheny (1981) noted that each of his three subjects, who attempted to maximize speech clarity for hearing-impaired listeners, appeared to have an individual strategy for altering normal conversational speech style. Although all three maximally distinguished vowel formant frequency characteristics, one subject's clarified speech was characterized by a lengthening of content words at the expense of function words, the second by lengthening of function words proportionately more than content words, and the third by lengthening of all words. Despite such differing strategies, the outcome was essentially the same, indicating that the speaker's objectives in producing "clear speech" could be satisfied independently of the means by which he or she "chose" to accomplish this task. Gay (1981) concluded from his evaluation of the literature that

> speakers do not control their rate of speech by either a single mechanism or along a single dimension. The fact that duration of segmental units, [and] the displacement . . . of articulatory movements . . . undergo nonlinear transformations during changes in speaking rate, precludes the operation of a single mechanism for rate control. (p. 158)

The results presented here would suggest that accuracy control is probably as complexly determined.

Given the present findings, it is relevant to review Delattre's (1969) hypothesis that it may be stress, and its secondary manifestations rather than duration, that determines vowel formant frequency characteristics. Stress is itself correlated with more intense amplitude of F_2 and with a relative increase of that formant frequency (Lieberman, 1967). That most of the significant shifts towards centralization of vowel formant frequencies in the present adult–adult function-word contexts were shifts along the F_2 axis may suggest that stress and its ancillary manifestations, taken as a complex, are better predictors of vowel formant frequency characteristics than is duration alone. This investigation examined only the durational properties of vowels and so does not provide evidence regarding the possible relationships among other features that convey stress, such as F_0 and intensity, and vowel formant characteristics. Although such variables would have been impossible to appraise during the informal play sessions recorded here, the results from this analysis of uncontrolled speech samples would suggest that formal appraisal of the relationship between the components of stress realization and articulatory target achievement is warranted.

For those addressing hearing-impaired or unsophisticated listeners in the educational or clinical setting, or for those manipulating speech production for intelligibility or recognition testing, the present data suggest that simply extending the duration of vowel segments does not necessarily guarantee that the vowel will be characterized by relatively canonical formant frequency properties, nor will shortened vowels necessarily be characterized by less canonical formant frequency characteristics than those seen in single-word productions. Additionally, although the women in this study slowed their overall speech rate to their children almost 25% (from 184 wpm to 138 wpm), such a global rate adjustment did not translate directly into longer segmental durations. This finding, together with the findings of other studies showing that the overall rate of speech cannot necessarily be mapped onto changes in the acoustic characteristics of phonemes, suggests that rate modification is too imprecise a parameter to induce changes in the intelligibility of the speech signal. Individual strategies, or more likely the intent of the speaker to emphasize certain aspects of the message, may be more important determinants of the clarity of the message than simple measures of duration or rate.

These concepts become more important as one moves away from the consideration of the bulk of the literature, which has examined rate and duration effects on the intelligibility and acoustic properties of the signal in the context of single-word production or the reading of text, and towards consideration of the way in which such parameters might interact with others in the context of conversational speech production. The hypotheses that a speaker entertains when asked to clarify individual words in an experimental paradigm, and how the speaker goes about this task, may differ significantly from goals and strategies for clarifying entire messages.

As a final observation, some researchers have begun to investigate the influence of speaking rate on the processing of language by aphasics (Pashek & Brookshire, 1982) or the learning of second languages (Hatch, 1983). Although slowed rate probably aids the processing abilities of a person with limited language skills, we might caution that for clinicians or researchers tempted to use rate manipulations to increase listener comprehension, perhaps it would be better if they considered their task to be that of speaking clearly rather than speaking slowly. The results of this investigation and prior research suggest that the second phenomenon is a by-product of the first, although the effects of rate itself on the clarity of speech are much less certain. More importantly, the speech register that reflects a speaker's desire to be clear conversationally is characterized by adjustments in the acoustic parameters which specify phoneme identity. Such adjustments do not appear to be related to simple changes in segmental duration in a straightforward manner.

ACKNOWLEDGMENTS

Preparation of this manuscript was supported by a University of Maryland General Research Board Summer Research Award. The support of the University of Maryland Computer Science Center is also acknowledged. The assistance of the following people at the Research Laboratory of Electronics at the Massachusetts Institute of Technology, which provided facility and materials support for this study, is gratefully acknowledged: Victor Zue, Keith North, Stephanie Seneff, and Lori Lamel. I am also indebted to Gary Weismer and a trio of anonymous reviewers for their insightful comments on an earlier draft of this work.

REFERENCES

BERNSTEIN-RATNER, N. (1984a). Patterns of vowel modification in mother–child speech. *Journal of Child Language, 11*, 557–578.

BERNSTEIN-RATNER, N. (1984b). Phonological rule usage in mother–child speech. *Journal of Phonetics, 12*, 245–254.

BERNSTEIN-RATNER, N., & PYE, C. (1984). Higher pitch in babytalk is *not* universal: Acoustic evidence from Quiche Mayan. *Journal of Child Language, 11*, 515–522.

BROAD, D., & WAKITA, H. (1977). A phonetic approach to automatic vowel recognition. In L. Bolc (Ed.), *Natural language communication with computers*. Berlin: Springer-Verlag.

DELATTRE, P. (1969). An acoustic and articulatory study of vowel reduction in four languages. *International Review of Applied Linguistics in Language Teaching, 7*, 295–325.

GAY, T. (1978). Effect of speaking rate on vowel formant movements. *Journal of the Acoustical Society of America, 63*, 223–230.

GAY, T. (1981). Mechanisms in the control of speech rate. *Phonetica, 38*, 148–158.

GAY, T., USHIJIMA, T., HIROSE, H., & COOPER, F. (1974). Effects of speaking rate on labial consonant-vowel articulation. *Journal of Phonetics, 2*, 47–63.

HATCH, E. (1983). *Psycholinguistics: A second language perspective*. Rowley, MA: Newbury House.

HOUSE, A., WILLIAMS, C., HECKER, M., & KRYTER, D. (1965). Articulation testing methods: Consonantal differentiation with a closed response set. *Journal of the Acoustical Society of America, 37*, 158–166.

KIRK, R. (1969). *Experimental design: Procedures for the behavioral sciences*. Belmont, CA: Wadsworth.

KLATT, D. (1975). Vowel lengthening is syntactically determined in connected discourse. *Journal of Phonetics, 3*, 129–140.

KOOPMANS-VAN BEINUM, F. (1980). *Vowel contrast reduction: An acoustic and perceptual study of Dutch vowels in various* *speech conditions*. Unpublished doctoral dissertation, University of Amsterdam.

LIEBERMAN, P. (1967). *Intonation, perception and language*. Cambridge, MA: MIT Press.

LINDBLOM, B. (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America, 35*, 1773–1781.

MALSHEEN, B. (1980). Two hypotheses for phonetic clarification in the speech of mothers to children. In G. Yeni-Komshian, J. Kavanagh, & C. Ferguson (Eds.), *Child phonology* (Vol. 2). New York: Academic Press.

MILLER, J. (1981). Effects of speaking rate on segmental distinctions. In P. Eimas & J. Miller (Eds.), *Perspectives on the study of speech*. Hillsdale, NJ: Lawrence Erlbaum Associates.

MONSEN, R., & ENGEBRETSON, A. (1983). The accuracy of formant frequency measurements: A comparison of spectrographic analysis and linear prediction. *Journal of Speech and Hearing Research, 26*, 89–97.

PASHEK, G., & BROOKSHIRE, R. (1982). Effects of rate of speech and linguistic stress on auditory paragraph comprehension of aphasic individuals. *Journal of Speech and Hearing Research, 25*, 377–382.

PETERSON, G., & BARNEY, H. (1954). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America, 24*, 175–184.

PICHENY, M. (1981). *Speaking clearly for the hard of hearing*. Unpublished doctoral dissertation, Massachusetts Institute of Technology.

TIFFANY, W. (1959). Nonrandom sources of variation in vowel quality. *Journal of Speech and Hearing Research, 2*, 305–317.

TOLLHURST, G. (1957). Effects of duration and articulation changes on intelligibility, word reception and listener preference. *Journal of Speech and Hearing Disorders, 22*, 328–334.

VERBRUGGE, R., & SHANKWEILER, D. (1977). Prosodic information for vowel identity. *Journal of the Acoustical Society of America, 61*, S39.

Received January 27, 1984
Accepted January 16, 1985

Requests for reprints should be sent to Nan Bernstein-Ratner, Department of Hearing and Speech Sciences, LeFrak Hall, University of Maryland, College Park, MD 20742.

APPENDIX A

Formant Frequency Characteristics of Monophthong Vowels in Adult–Adult (A–A) and Adult–Child (A–C) Speech

Vowel	*A–A Content–Function*	*A–C Content–Function*	*A–A Content, A–C Content*	*A–A Function, A–C Function*
/æ/F_1	793 756	815 829	793 815	756 829
	$t(56) = 0.8231, p \le .4140$	$t(55) = -0.3053, p \le .7613$	$t(51) = -0.4846, p \le .6300$	$t(60) = -1.5953, p \le .1159$
F_2	1981 1911	2010 2019	1981 2010	1911 2019
	$t(56) = 1.5204, p \le .1340$	$t(55) = -0.1590, p \le .8742$	$t(51) = -0.6219, p \le .5368$	$t(60) = -2.1465, p \le .0359$
/ɛ/F_1	633 602	636 683	633 636	602 683
	$t(15) = 0.4452, p \le .6625$	$t(11) = -0.5867, p \le .5810$	$t(19) = -0.0691, p \le .9456$	$t(7) = -0.7725, p \le .4651$
F_2	2135 1761	2200 2239	2135 2200	1761 2239
	$t(15) = 4.0659, \underline{p \le .0010}$	$t(11) = -0.2595, p \le .8001$	$t(19) = -0.7185, p \le .4812$	$t(7) = -3.5745, p \le .0090$
/ɪ/F_1	469 463	471 485	469 471	463 485
	$t(67) = 0.2498, p \le .8035$	$t(63) = -0.5388, p \le .5919$	$t(55) = -0.0969, p \le .9231$	$t(75) = -0.9083, p \le .3666$
F_2	2253 2038	2260 2192	2253 2260	2038 2192
	$t(67) = 2.6824, p \le .0064$	$t(63) = 1.0458, p \le .2996$	$t(55) = -0.1044, p \le .9173$	$t(75) = -2.0645, p \le .0424$
/ʊ/F_1	520 534	530 478	520 530	534 478
	$t(38) = -0.2245, p \le .8236$	$t(39) = 0.9416, p \le .3522$	$t(72) = -0.4171, p \le .6779$	$t(5) = 0.6449, p \le .5474$
F_2	1164 1539	1207 1230	1164 1207	1539 1230
	$t(38) = -3.2364, \underline{p \le .0025}$	$t(39) = -0.1710, p \le .8651$	$t(72) = -0.8577, p \le .3939$	$t(5) = 1.2704, p \le .2599$
/a/F_1	796 820	768 799	796 768	820 799
	$t(46) = -0.8653, p \le .3914$	$t(32) = -0.7475, p \le .4602$	$t(43) = 1.0834, p \le .2847$	$t(35) = 0.5646, p \le .5759$
F_2	1358 1481	1487 1462	1358 1487	1481 1462
	$t(46) = -2.1340, p \le .0382$	$t(32) = 0.2645, p \le .7931$	$t(43) = -1.8194, p \le .0758$	$t(35) = 0.2440, p \le .8087$
/e/F_1	618 564	652 564	618 652	564 564
	$t(44) = 1.2926, p \le .2029$	$t(27) = 1.6620, p \le .1081$	$t(59) = -1.1509, p \le .2544$	$t(12) = -0.0087, p \le .9932$
F_2	2190 2040	2351 2394	2190 2351	2040 2394
	$t(44) = 1.6761, p \le .1008$	$t(27) = -0.4769, p \le .6373$	$t(59) = -2.6866, p \le .0094$	$t(12) = -4.2847, \underline{p \le .0011}$
/i/F_1	414 413	398 391	414 398	413 391
	$t(50) = 0.0613, p \le .9514$	$t(41) = 0.2817, p \le .7796$	$t(41) = 0.6103, p \le .5450$	$t(50) = 0.9615, p \le .3409$
F_2	2391 2437	2495 2491	2391 2495	2437 2491
	$t(50) = -0.8760, p \le .3852$	$t(41) = 0.0888, p \le .9297$	$t(41) = -1.8317, p \le .0743$	$t(50) = -1.2153, p \le .2300$
/u/F_1	424 430	412 369	424 412	430 369
	$t(36) = -0.3155, p \le .7542$	$t(24) = 2.1581, p \le .0411$	$t(35) = 0.6001, p \le .5523$	$t(25) = 3.1616, p \le .0041$
F_2	1076 1089	1073 990	1076 1073	1089 990
	$t(36) = -0.1961, p \le .8456$	$t(24) = 1.0918, p \le .2857$	$t(35) = 0.0406, p \le .9679$	$t(25) = 1.1738, p \le .2515$

Note. The underlined values are significant.

APPENDIX B

Vowel Durational Characteristics of Monophthong Vowels in Adult–Adult (A–A) and Adult–Child (A–C) Speech

Vowel	*A–A Content–Function*			*A–C Content–Function*			*A–A Content, A–C Content*			*A–A Function, A–C Function*		
/æ/	150	105	$t(56) = 2.8782, p \le .0057$	142	105	$t(55) = 1.9432, p \le .0571$	150	142	$t(51) = 0.4256, p \le .6722$	105	105	$t(60) = 0.0171, p \le .9864$
/ɛ/	113	81	$t(15) = 0.9798, p \le .3427$	132	72	$t(11) = 1.0883, p \le .2997$	113	132	$t(19) = -0.5332, p \le .6001$	81	72	$t(7) = 0.4070, p \le .6962$
/ɪ/	103	86	$t(67) = 2.0664, p \le .0425$	98	92	$t(63) = 0.4716, p \le .6388$	103	98	$t(55) = 0.4044, p \le .6875$	86	92	$t(80) = -0.7825, p \le .4362$
/ʊ/	88	59	$t(38) = 1.0491, p \le .3008$	83	93	$t(39) = -0.5320, p \le .5977$	88	83	$t(72) = 0.5705, p \le .5701$	59	93	$t(5) = -1.0153, p \le .3566$
/a/	155	126	$t(46) = 1.8776, p \le .0668$	142	92	$t(32) = 3.1029, p \le .0046$	155	142	$t(43) = 0.7015, p \le .4868$	126	92	$t(35) = 2.6598, p \le .0117$
/e/	146	108	$t(44) = 1.8713, p \le .0680$	156	128	$t(27) = 0.9037, p \le .3742$	146	156	$t(59) = -2.3375, p \le .0228$	108	128	$t(12) = -0.9291, p \le .3712$
/i/	125	131	$t(50) = -0.3035, p \le .7628$	135	111	$t(41) = 1.4228, p \le .1624$	125	135	$t(41) = -0.4675, p \le .6427$	131	111	$t(50) = 1.2078, p \le .2328$
/u/	157	79	$t(36) = 5.4115, p \le .0001$	134	157	$t(24) = -0.7046, p \le .4879$	157	134	$t(35) = 1.1488, p \le .2584$	79	157	$t(25) = -3.0554, p \le .0053$

Note. The underlined values are significant.

APPENDIX C

Formant Frequency and Durational Characteristics of Diphthong Vowels in Adult–Adult (A–A) and Adult–Child (A–C) Speech

Vowel	*Characteristic*	*A–A*	*A–C*	
/ai/	F_1 onset	815	766	$t(50) = 1.1373, p \le .2608$
	F_1 offglide	501	462	$t(50) = 0.8855, p \le .3803$
	F_2 onset	1426	1425	$t(50) = 0.0054, p \le .9957$
	F_2 offglide	2183	2345	$t(50) = -1.8998, p \le .0635$
	duration	121	127	$t(50) = 0.6161, p \le .5406$
/au/	F_1 onset	855	880	$t(65) = -0.8671, p \le .3889$
	F_1 offglide	553	511	$t(65) = 1.0647, p \le .2910$
	F_2 onset	1730	1851	$t(65) = -2.0187, p \le .0475$
	F_2 offglide	1109	1045	$t(65) = 1.0970, p \le .2767$
	duration	171	175	$t(65) = -0.3503, p \le .7272$
/ou/	F_1 onset	690	676	$t(78) = 0.4272, p \le .6704$
	F_1 offglide	476	456	$t(78) = 0.7018, p \le .4852$
	F_2 onset	1463	1491	$t(78) = -0.4418, p \le .6599$
	F_2 offglide	1023	1028	$t(78) = -0.0819, p \le .9350$
	duration	147	148	$t(78) = -0.0447, p \le .9644$
/ɔi/	F_1 onset	565	503	$t(17) = 0.8522, p \le .4059$
	F_1 offglide	446	392	$t(17) = 0.7973, p \le .1901$
	F_2 onset	1049	965	$t(17) = 0.7973, p \le .4363$
	F_2 offglide	2060	2381	$t(17) = -1.5459, p \le .1406$
	duration	273	287	$t(17) = -0.2973, p \le .7698$