WORD- AND PHRASE-LEVEL ASPECTS OF VOWEL REDUCTION IN ITALIAN.

E. Farnetani and M. Vayra

Centro di Fonetica del CNR, Padova, Italy
Scuola Normale Superiore, Pisa, Italy

ABSTRACT

The present study investigates some of the sources of vowel reduction in Italian and the relationship between reduction and acoustic duration. The effects of stress, position within a word, and position within an utterance on the spatial and temporal characteristics of vowel /a/ were quantified and compared. The effects of position within an utterance (initial vs final) were investigated to verify the hypothesis of a progressive early-to-late reduction. The results indicate that phrase-level reduction is unsystematic, (it was observed only in one of our two subjects), and does not appear to be progressive (it is confined to the very edges of the utterance). The most relevant and systematic position effects are instead the lengthening and opening of unstressed vowels in utterance final position. Two findings emerge from the various patterns of interaction between duration and first formant frequency: a) speakers can control the two variables independently and such a control appears to be addressed to the preservation of stress contrast; b) the extent of such control seems to depend on the more or less elaborated speech style that characterizes different speakers.

1. INTRODUCTION

This study concerns articulatory and spectral reduction of vowels in Italian, the former defined as a decrease in the magnitude of gestural displacement, the latter as an increased amount of centralization within the acoustic vowel space.

Two issues are the focus of the present work: the analysis of possible sources of vowel reduction, and the assessment of the relationship between reduction and acoustic duration.

Evidence of spatiotemporal reduction of unstressed vowels in Italian has emerged from a number of studies ([1], [2]). A recent experiment on tongue dorsum and jaw movements [3] showed that unstressed vowel /a/ is subject to a high degree of reduction, manifested as a decreased displacement of the jaw from the rest position (see also [4]). Other studies on Italian have shown that duration is affected by position and by the number of preceding and of following segments within and across a word [5]. Such data are compatible with the timing model proposed by Lindblom & Rapp [6]. No studies, however, have been carried out to establish whether or not temporal compressions are paralleled by spatial reductions, which should be expected to occur according to the duration-dependent undershoot hypothesis put for by Lindblom [7].

Results of a recent investigation [4] suggest that another source of reduction can be at play at phrase-level: a progressive early-to-late reduction, inferred from a monotonic decrease in first formant frequency and in the amplitude of jaw opening of stressed /a/ along an utterance. If this kind of phrase-level reduction occurs, to what extent will duration be compatible with a duration-dependent undershoot model, which predicts an increased degree of reduction in shorter segments? Investigations concerning the relationship between duration and reduction show that the two variables can be controlled independently. Engstrand [8] showed that in fast speech stressed vowels are not more reduced than their longer counterparts in slow speech. Similarly, Nord [9] showed that finally lengthened unstressed vowels tend to be more centralized than non-final stressed vowels of the same duration.

The present speech material has been constructed in such a way that all the above mentioned potential sources of reduction can be investigated and their effects quantified and compared, while an assessement of the relationship between duration and reduction has made it possible to identify the conditions under which the two variables appear to be dissociated.

2. METHOD

The corpus consists of trisyllabic nonsense words of the type /CVCVCV/, where C = /t/ and V = /a/, or /i/, or /u/ in symmetric sequences, with stress on initial , medial and final position. The key words were repeated five times in three contexts: in isolation, in sentence initial position *(Ugo.... della Torre parti' per la Francia),* and in sentence final position *(Parti' per la Francia col marchese Ugo ....).* Subjects were one female (S1) and one male (S2) Northern speakers of Standard Italian. We simultaneously collected acoustic and electro palatographic (EPG) data. We shall report here the subset of results relative to /a/. For low vowels, a decrease in first formant frequency and an increase in amount of linguopalatal contact in the back region indicate higher tongue/jaw position. They were used as indices of reduction. The durations of vowels were defined as the intervals of periodicity within each syllable; EPG and first formant values were measured at vowel mid points. It must be reminded that EPG provides only partial information on vowel configuration, thus, especially for vowel /a/ we have relied more heavily on F1 than on EPG. The following variables were tested: stress, syllable position within the word, position of the word within the utterance. Series of ANOVAs and t-tests were used for statistical analyses. In the description of the results, we shall refer to differences with a significance level no less than 97.50%.

3. RESULTS AND COMMENTS

3.1 Reduction

In both subjects stressed vowels have higher F1 and less linguo-palatal contact than unstressed vowels. For S1, F1 decreases by 28% in unstressed vowels (average values: 974 vs. 700 Hz), for S2 it decreases by 18% (average values: 669 vs. 548 Hz); for S1 the EPG contact area decreases during stressesd vowels by 80% (average contact: 4.39 vs 0.89); for S2 it decreases by 53% (average contact: 5.68 vs 2.67).

Table 1 shows the mean F1 values.

TABLE 1.

Mean values of F1 (Hz). W1, W2, W3 refer to words with stress on the first, on the second and on the third syllable. Numbers 1, 2, and 3 refer to the context in which the test word was produced: isolated, sentence initial, and sentence final, respectively.

CONTEXT	1	2	3
S1			
W1	**980** 666 855	**990** 661 526	**983** 619 820
W2	850 **983** 882	554 **988** 514	626 **1000** 873
W3	956 711 **959**	559 644 **962**	631 656 **920**
S2			
W1	**654** 554 654	**685** 522 479	**634** 524 591
W2	564 **739** 655	459 **606** 528	496 **654** 617
W3	566 539 **738**	504 530 **690**	479 536 **674**

Globally both stressed and unstressed vowels are affected by position within word or sentence: the higher F1 of unstressed vowels in final position (contexts 1 and 3) indicates that they tend to open, while stressed vowels in final position tend to be more reduced than non-final vowels in S1 (contexts 2 and 3, although only in context 3 the difference reaches the significance level $p<0.02$), whilst they do not change significantly in S2.

As for the stressed vowels produced by S1, the fact that only those in utterance final position show some significant differences with earlier vowels, and, as shown in Table 1, tend to be more reduced than any other stressed vowels earlier in the sentence, indicates that such reduction is a phrase-level rather than a word-level phenomenon; at the same time it suggests that weakening of stressed vowels is confined to the very last syllable of the utterance. Thus, the present data are only in partial agreement with the hypothesis of a monotonic early-to-late reduction. As for unstressed vowels, the fact they do not tend to open in final position of utterance-initial words, which are not phrase final, suggests that also the decrease in reduction of unstressed vowels in final position is to be regarded as a phrase-level phenomenon. This is corroborated by another observation relative to S1: in isolated words also the unstressed vowels

of initial syllables tend to open (see Table 1), and only in this context the initial syllables of the word are also in phrase- (and utterance-) initial position. Thus, the patterns of reduction for unstressed vowels can be interpreted in very simple terms as tendencies to alterate the degree of vowel height at prosodic boundaries, or, better, in absolute final position for both speakers, and also in absolute initial position for S1. All the other unstressed vowels appear to be equally reduced whatever is their position.

The EPG data do not show the trend observed in F1 for the stressed vowels, but capture the two degrees of reduction observed for unstressed vowels: for both speakers the amount of EPG contact in vowels adjacent to prosodic boundaries is about 23% less than the contact in the more reduced vowels.

3.2 Duration and reduction

The global data show, as expected, that duration and F1 are highly correlated. The correlation coefficients are r(133) = 0.783 for S1, and r(130) = 0.774 for S2. Duration is also correlated with the amount of EPG contact with r(130) = -0.704 for S1 and r(130) = -0.654 for S2. We shall examine in detail the relationship between duration and F1, since duration accounts for a larger proportion of variance of F1 than of EPG for both subjects.

TABLE 2.

Mean vowel durations (ms). Captions as in Tab.1

CONTEXT	1	2	3
S1			
W1	**140** 59 134	127 54 58	131 57 136
W2	68 **157** 142	55 **126** 47	43 **150** 106
W3	65 59 **188**	52 61 **116**	60 53 **127**
S2			
W1	**103** 60 114	**98** 56 71	**100** 58 68
W2	62 **142** 116	41 **109** 41	49 **117** 71
W3	61 53 **144**	53 64 **95**	53 65 **102**

The table shows, as expected, longer durations for stressed than for unstressed vowels. The comparison with Table 1, indicates that the differences in duration between stressed and unstressed vowels are much more often neutralized by position effects than the differences in F1, and that the relations between duration and F1 differ in the two subjects. For S1, in isolated words, we observe that unstressed vowels undergo final lengthening, which, as seen before, is accompanied by an increase in F1; instead the unstressed vowels in the initial syllables are not significantly longer than medial vowels, in spite of their remarkably high F1 values (cf. Table 1). A second important discrepancy between duration and F1 is observed in the utterance final unstressed vowels: final lengthening makes them as long as the preceding stressed vowels (see Table 2, Context3, W1) or as long as stressed vowels occupying the same position (see Context3 W1 vs W3). They have, however, significantly lower F1 values than the stressed vowels of the same duration even though such values are higher than those of the shorter unstressed vowels (cf. Table 1). Taken together the data indicate that the speaker has allowed the unstressed final vowels to lengthen and to open, but has prevented them from opening as much as the stressed vowels of the same duration. The dissociation between duration and F1 in the first syllables of isolated words could instead be viewed as an effect of a greater articulatory force characterizing the very beginning of an utterance. As for stressed vowels, their variations in duration are not reflected in F1, which appears to be rather stable: only in utterance final position the shortening of the stressed vowel is associated with a decrease in F1. This suggests that the speaker could vary stressed vowel duration without varying their height.

In Figure 1 the mean values of F1 are plotted against the mean durations; the figure also shows the regression line obtained from the row data.

As for S1, (on the left), the figure shows that stressed an unstressd vowels are always distinguished along one or the other dimension. Stressed vowels vary in duration much more than in F1. The unstressed vowels tend to fall into three groups: short (utterance-initial) ureduced vowels, long (utterance-final) moderately reduced vowels, and short highly reduced vowels (those not adjacent to boundaries). Altogether the data indicate that S1 exerted independent control over duration and gestural

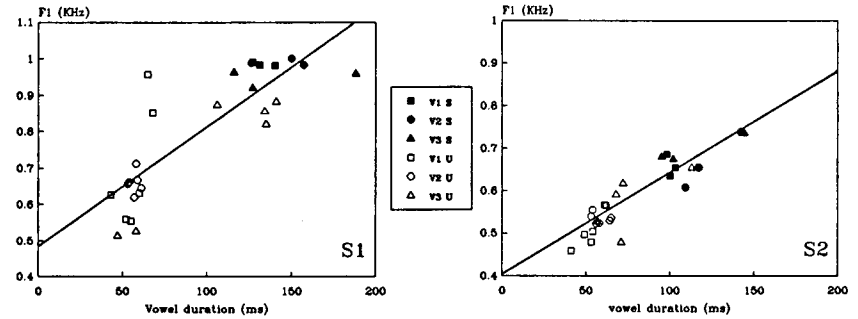

FIGURE 1. F1 plotted against duration; 1, 2, 3: V position within word.

amplitude in both stressed and unstressed vowels.

As for S2, in isolated words the duration patterns reflect quite well the F1 patterns, with the consequence that unstressed final vowels, which undergo a remarkable final lengthening (and an increase in F1), are no longer distinguished from stressed vowels in initial position (cf. Figure 1 and Tables). In embedded words the very short durations of unstressed vowels parallel their very low F1 values. Instead, stressed vowels, athough characterized by longer durations in word medial positions, have F1 frequencies tendentially lower than those of vowels in initial and final positions. This indicates that S2 varied the movement velocity in producing stressed vowels in different word position. The data suggest that S2 exerted an independent control over duration and gestural amplitude mostly in the production of stressed vowels, thus, much less extensively and systematically than S1.

5. CONCLUSION

The overall F1 and EPG data on vowel /a/ indicate that 1) stress seems to be the major factor in vowel reduction; 2) adjacency to prosodic boundary has more systematic influence on unstressed than on stressed vowels; 3) in S1 there is clear evidence of reduction of stressed vowels in utterance final position. As for the intersubject differences in the interactions between F1 and duration, we ascribe the more extensive independent control exhibited by S1 to her more elaborated speech style. Evidence that S2 used a more casual speech style than S1 emerges from the combination of his lower F1 values, larger areas of EPG contact, shorter durations, and higher coefficients of variation.

The finding that this speaker exerted independent control over duration and F1 less extensively than S1 fits quite well in the global speech style picture we are proposing. The comparison between S1 and S2 data in Fig.1 bears witness to our interpretation: in S2 we note, together with reduced distances between the two stress categories, less extensive deviations from the regression line and more data points below than above it.

6. REFERENCES

[8] Engstrand, O. (1988), "Articulatory correlates of stress and speaking rate in Swedish VCV utterances", *J.A.S.A., 83*, 5, 1863-1875.

[1] Farnetani, E. & S. Kori (1981), "Italian lexical stress in connected speech". In *Proc. 4th F.A.S.E. Symposium, 1*, Roma: Edizioni Scientifiche Associate, 57-61.

[3] Farnetani, E. & A. Faber (1991), "Vowel production in isolated words and in connected speech: An investigation of the linguo-mandibular subsystem", in press.

[7] Lindblom, B. (1963), "A spectrographic study of vowel reduction", *J.A.S.A., 35*, 1773-1785.

[6] Lindblom, B. & K. Rapp (1973), "Some temporal regularities of spoken Swedish". *Publ. no. 21, University of Stockholm*: Institute of Linguistics.

[9] Nord, L. (1986), "Acoustic studies of vowel reduction in Swedish", *STL-QPRS, 4*, 19-36.

[5] Vayra, M., C. Avesani & C. Fowler (1984), "Patterns of temporal compression in spoken Italian". In *Proc. Xth ICPhS*, Dordrecht, Holland: Foris Publ., 541-546.

[2] Vayra, M. & C. Fowler (1987), "The word-level interplay of stress, coarticulation, vowel height and vowel position in Italian". *Proc. XIth ICPhS, 4*, Tallinn, Estonia: Academy of Sciences of the Estonian S.S.R., 24-27.

[4] Vayra, M. & C. Fowler, "Declination of supralaryngeal gestures in spoken Italian", submitted.