The phonetic word: the articulation of stress and boundaries in Italian

Edda Farnetani
Istituto di Fonetica del CNR
Padova (Italy)

1.Introduction

A spoken word produced in its canonical form can be defined as a phonetic entity with its specific segmental and suprasegmental structure. The suprasegmental component is expressed by two sets of acoustic/perceptual attributes: temporal/melodic properties at the word's edges, signaling initial and final boundaries, and, for polysyllabic words, prominence contrast between syllables. In free-stress languages prominence at the word level is determined by the position of lexical stress.
In normal speech communication, the word, unless uttered in isolation or produced in clear speech style, tends to lose its canonical properties and to undergo a number of changes both in the segmental and the suprasegmental properties. Such changes usually referred to as connected speech processes (see Kohler, 1990 for a review), range from segmental reductions, weakening of stress contrast, assimilations, to feature and segment deletion, cancellation of boundary signals and of lexical prominence. In this case the word disappears as a phonetic unit. A great number of studies show that speaking rate and speech style are two important factors affecting the articulation of a segment (see Perkell, 1977, p.352 for a review of the changes in the articulatory kinematics as a function of these two factors); linguistic factors such as prosodic/syntactic structure have traditionally been investigated in terms of acoustic melodic and/or durational effects. Only recent research has been directed to the analysis and the modeling of the changes in the articulatory configuration and in the kinematics of gestures as a function of prosodic structure (see below, 2.1 for a review of recent studies on English and French).
This experiment, which expands previous research on this topic (Farnetani and Vayra, 1996) deals with the supraglottal articulation of lexical stress and of word boundaries, and tests the effects of three different speech contexts: isolated words, sentence final words, and words embedded in a three-component syntactic phrase. Thus it compares citation forms with words in two syntactic/prosodic positions, phrase/sentence final position, and phrase-medial position. The 1996 study dealt with the spatial effects of stress and boundaries, the present paper is concerned with both the spatial and the temporal/dynamic aspects of stress and boundary production.
One purpose if this study is to assess how the articulation of vowels and consonants in CV syllables signals prominence and initial and final boundaries, and whether and how the articulation changes as a function the three prosodic contexts; a second purpose is to get more knowledge of the dynamics of stress and boundary production, in the light of current phonetic models of their realization.

1.1. Previous studies on lexical stress and boundaries in Italian.

As for boundaries, the fact itself that in continuous speech words combine together to form larger prosodic units implies that at the lexical level boundary signals are weaker than at any higher level. The question is then: are word level boundaries preserved in items embedded in larger prosodic units, and if so, to what extent are they weakened? Or are they totally canceled?. As for lexical stress, we can ask whether the syllabic prominence associated with stress weakens in embedded words. A previous acoustic study on boundaries in Italian (F0 and duration) indicated that within two-word phrases, both final boundaries of W1 and initial boundaries of W2 disappear, while syllabic prominence is preserved in both words (Farnetani, 1989). Other acoustic studies on lexical stress and rhythmic structure (Farnetani and Kori, 1983, 1990) indicate that in phrase-non-final bisyllabic words lexical prominence is preserved, but that stressed vowels undergo a progressive durational shortening as the interstress interval between the target word and the following one decreases. The 1983 study shows that word level prominence is totally canceled only under the condition of stress clash.
The overall data clearly indicate that in continuous speech word level prominence resists reduction and deletion more than boundary signals, and therefore becomes itself a cue to word parsing. This idea is

supported by a number of ASR studies (Lea, 1980), where the function of lexical prominence for detecting lexical items in continuous speech has been tested for English: the results show that prominent syllables are not only anchor points for phonetic recognition, but also a guide to the listeners for word detection, even if they do not, *per se,* delimit words from the neighbours.

1.2. Current models of stress and boundaries production

Three models of the production of prominence have recently been proposed: the "Jaw expansion" model of Macchi (1985), the "Sonority" model of Beckman, Edwards and Flechter (1992), the "Hyperarticulation" model, based on Lindblom's Hyper/Hypospeech theory (1990), and proposed by de Jong (1995) to account for the articulation of prominence. Although quite different, all three of them seem to account for the main acoustic effects of prominence observed in English and in other languages: longer durations, more extreme vowel-formant frequencies, and higher intensity.

The jaw expansion model attributes the acoustic effects of stress to an expansion in space and time of jaw movement cycles, i.e. a greater and longer opening for vowels, and a more extreme and longer closing for consonants.

According to the sonority model, framed within the task-dynamic model of speech production (Browman and Goldstein (1989), it is the whole vocal tract to expand its cycles in space and time, thus both jaw, tongue, and lip movements increase in magnitude and duration in the production of prominence. Expansion is brought about by decreasing the temporal overlap between adjacent gestures, and aims at increasing the contrast along the scale of the sonority feature. Thus, making a syllable prominent means to maximize the sonority contrast between C and V of the syllable.

The H&H theory, based on the movement-target model of speech production, assumes that speech varies along a dimension from hypo- to hyper-articulation: at one end, hyperarticulation i.e. target overshoot aims at maximizing the segment phonetic distinctiveness, and is proper of clear speech style, at the other extreme, hypoarticulation i.e. target undershoot, aims at maximizing motor economy and characterizes colloquial, casual speech.

de Jong (1995) tested the three models, by analyzing the movements of the jaw, the tongue and the lips in monosyllabic words with nuclear accent (specifically in narrow focus condition), vs words with prenuclear accent, vs unaccented words. The observed tongue and (lower) lip positions were decomposed into independent jaw, tongue and lip components, which made it possible to infer the active movements of the tongue and lips. The results show that for nuclear /U/ the tongue articulator is indeed active in the production of prominence (contrary to the jaw expansion model), but its movements do not only expand along the high-low dimension, but also along the front-back dimension, suggesting a more posterior target specification; and the upper lips enhance their protrusion, as compared to the prenuclear and the unaccented /U/. Moreover the kinematics of gestures shows that the movement expansion in nuclear words is associated more strongly with an increase in velocity than in duration. According to de Jong, the hyper/hypospeech model is the only one that can account for all the findings, therefore nuclear prominence has to be considered an instance of localized hyperarticulation.

It must be noted that study by de Jong compares emphatic with prefocal accent with no accent, while the material used by Edwards, Beckman and Fletcher (1991); Beckman et al. (1992) compares phrase-final accented words with phrase nonfinal unaccented words, hence, the levels of prominence specification compared in the two studies are not the same.

As for final prosodic boundaries, the experiment by Edwards et al.(1991) proposes two alternative production strategies for the articulation of final boundaries: slowing down the speech tempo (by changing gestural stiffness), decreasing gestural overlap (by changing intergestural phasing). Both strategies bring about final syllable lengthening, but the second has also the effect of expanding the articulatory gestures and is applied in most cases for unstressed final syllables.

For the articulation of initial boundaries the EPG data for English by Keating (1995; in press) indicate that the tongue-to-palate contact in alveolars is always greater in word-initial than in word-final position and that in sentence initial words the contact of initial Cs is greater than in words in other positions. In line with de Jong's hyperarticulation hypothesis, Keating proposes that consonant strengthening in initial boundaries is an instance of hyperarticulation. The EPG data by Fougeron and Keating (1996) for French refer to tongue-to-palate contact for /n/ and /t/ production, and show that, at the word level, initial Cs have more contact than word medial Cs, and that the contact in initial consonants increases gradually as the prosodic level becomes higher, the greatest contact being associated with utterance initial position.

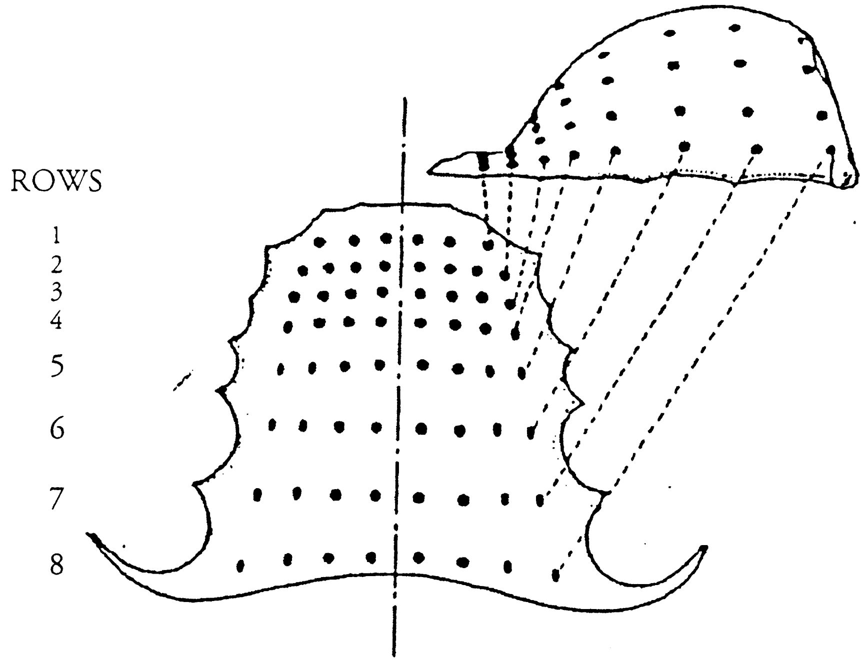

Fig. 1: Artificial palate: the electrodes are distribuited in eight rows from front to back.

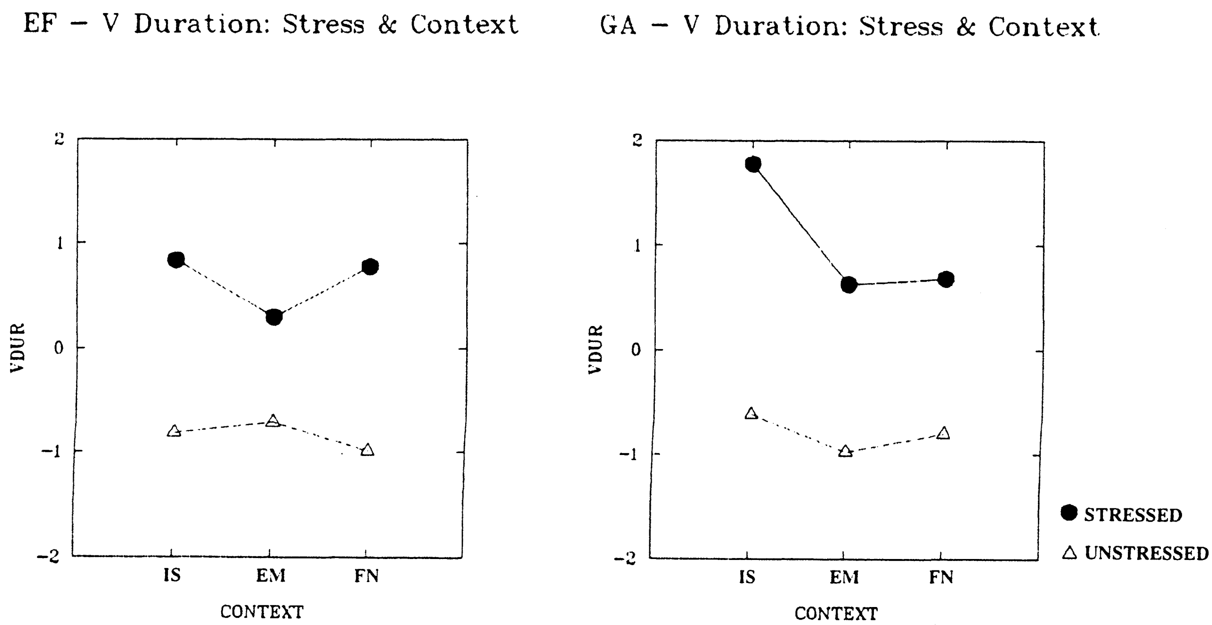

Fig. 2: Vowel durations, averaged across /i/, /a/, /u/, as a function of stress and context

2. Methods

2.1 Speech Material

The patterns of tongue movements and configurations in the production of stress and boundaries were investigated by means of EPG in trisyllabic CVCVCV words with varying lexical stress position. where C=/t/ and V= /a/ or /i/ or /u/ in symmetric sequences. The target words are the reiterant versions of the proper names *Giacomo, Nicola, Niccolò* (stress on syll.1, 2, and 3 respectively). The words were produced in isolation (IS), in sentence final position (FN) (*Partì per la Francia col marchese Ugo...*), and embedded in a three-constituent phrase in sentence initial position (EM) (*Ugo...della Torre partì per la Francia)*. The utterances were produced with no emphatic stress on any of the words: the isolated words and the sentence final words were produced with the expected sentence accent, and the embedded words with no accent or a weak accent (which preceded the phrasal accent of the phrase final word).

The material was repeated 5 to 6 times at the most comfortable reading rate by two Northern speakers of Standard Italian (EF and GA).

The use of trisyllabic words has allowed independent analysis of stress and boundaries: the effects of stress were analyzed paradigmatically in word-medial syllables, those of boundaries were analyzed by comparing word-medial syllables with word-initial and word-final ones (with both paradigmatic and syntagmatic comparisons).

The data were collected by means of the Reading EPG system which simultaneously records EPG and acoustic data at 10 ms rate. See, in Fig. 1 the distribution of electrodes along eight rows ordered from front to back.

2.2. Measurements and analyses

For the consonant the data were taken in the EPG frame of maximum front contact in the area extending from row 1 to row 5, which covers the alveolar and the prepalatal region, and the parameter is CFRONT, i.e. the percent electrode activation in this region. This large area was chosen in order to avoid ceiling effects in the articulation of consonants in prominent syllables.

For the vowels the measures were taken in the frame of maximum contact for /i/ and in that of minimum contact for /u/ and /a/. The vowel parameters are: PI (posteriority index), CI (centrality index).

PI indicates the variations in contact along the front/back dimension and is expressed by the row number where the frontmost contact occurs: PI ranges from 1, when the first row is contacted, to 9, when no contact occurs in any of the rows (which sometimes occurs for stressed /a/). Normally, the frontmost contact in the production of vowels is a contact at the sides of the palate. The right/left asymmetries were accounted for by adding to the row number a fixed value of 0.50 when the contact was on one side only.

CI, taken in the palatal and postpalatal region (Row 6 to 8) is a measure of the extent of tongue contact from the sides to the center of the palate, and is expressed by the number of activated electrodes in the row of maximum contact. CI ranges from 0 (no contact) to 8 (when all the eight electrodes of the row are contacted), and is corrected by subtracting a fixed value of 0.33 or 0.66, when one or two of the back rows, respectively, exhibit less central contact than the one with maximun contact. As can be seen in the palate profile of Fig.1, in this area an increase in centrality indicates a movement of the tongue body to a higher position.

Temporal parameters are: C closure duration, acoustic vowel duration.

In order to shed light on how the various changes are brought about in the CV syllable, the following measures were taken for /ta/ syllables: 1) MaxC (n. activated electrodes for /t/ in the last frame of max closure; 2) Vmin (n. of free electrodes in the first frame of minimum contact for /a/); 3) duration of the interval between CMax and Vmin; 4) difference in contact, and 5) mean and max rate of contact change between CMax and Vmin. The last three measures should reflect the amount of tongue displacement, its duration and rate in the C-to-V gestures The present paper will refer only to the very preliminary results of this analysis.

Comparisons among parameters were carried out on standardized values. The significance level of statistical analysis (ANOVA) was set at $p<0.02$.

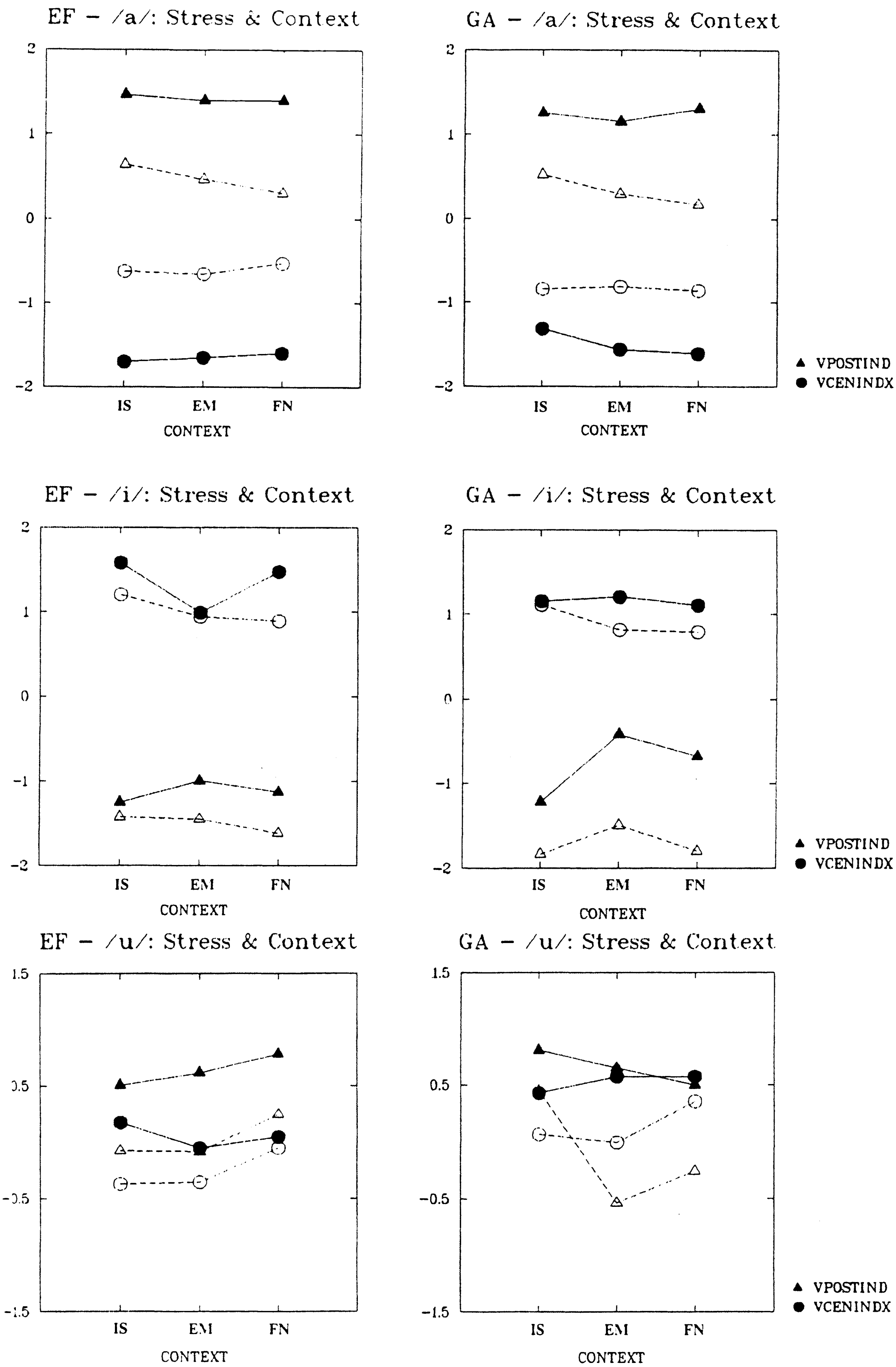

Fig. 3: Mean values of PI and CI for /a/, /i/, and /u/, as a function of stress and context

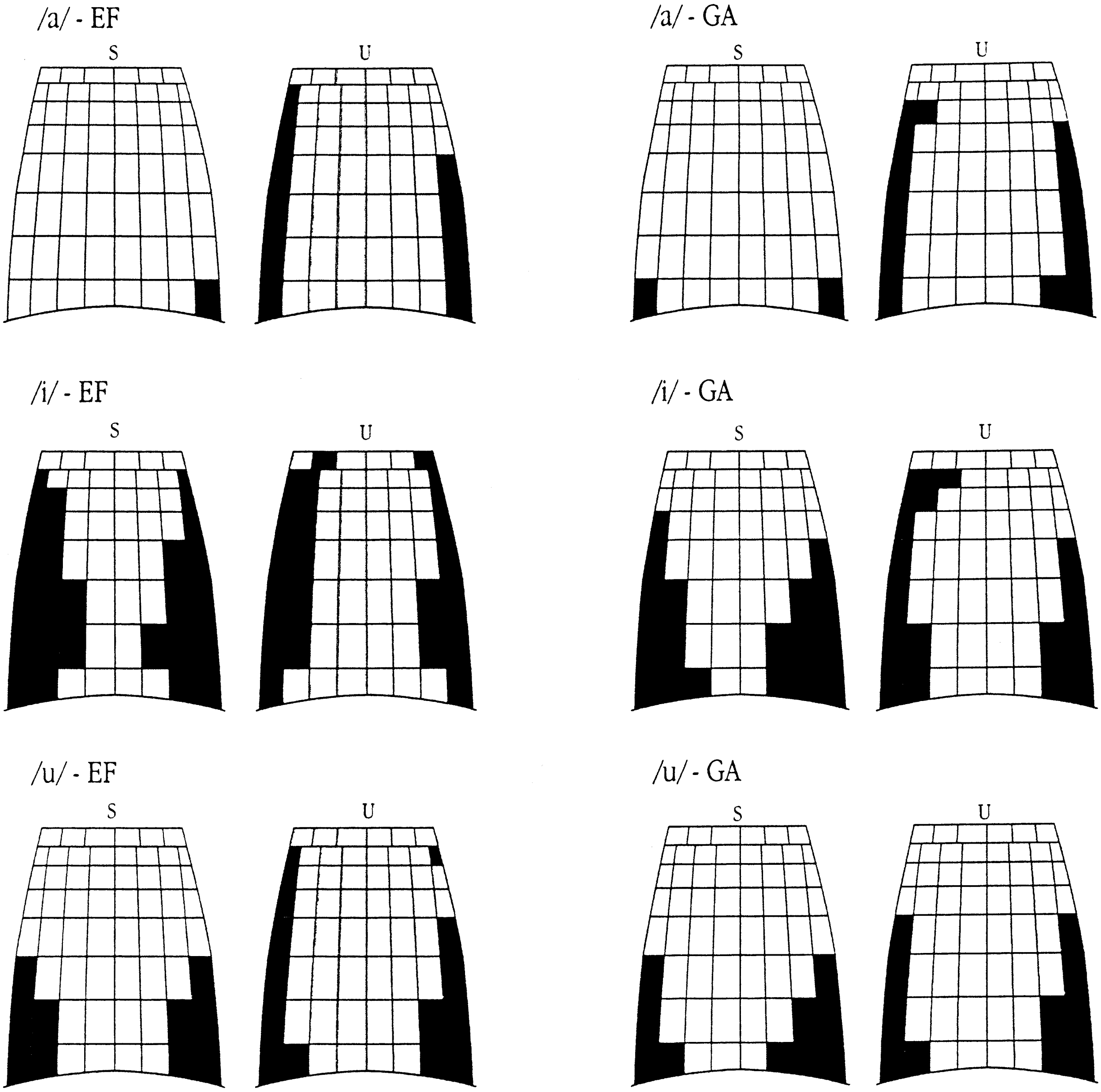

Fig. 4: Typical palatographic configurations of stressed and unstressed vowels.

3. Results

3.1. Prominence

3.1.1. Prominence in vowels

The results of comparisons between word-medial stressed and unstressed syllables, across the different prosodic conditions indicate that stress has significant effects on both the duration and the spatial configuration of all three vowels. The direction of the changes is the same for the two subjects. The data are illustrated in Fig.2, which shows the acoustic durations (averaged across the vowels), and Fig.3, which shows the spatial parameters for each vowel, in each prosodic context.

As for duration, stressed vowels are always significantly longer than unstressed ones, as expected (see Fig.2). The average durations are: ms 113.67 (S), ms 56.05 (U). Also the prosodic context has systematic effects on vowel duration for both subjects. Interactions between context and stress indicate that only stressed vowels change significantly as a function of context. As illustrated in Fig.2, for EF the stressed Vs are shorter in embedded words than in the other two conditions (ms 126.71 (IS), ms 105,20 (EM), ms 124.36 (FN), while for GA, the stressed Vs of embedded and final words are both shorter than the stressed Vs of isolated words (ms 130.60 (IS), ms 96.00 (EM), ms 97.76 (FN).

As for the spatial configurations (see Fig.3), stressed /a/ is produced with a more posterior, a more peripheral, and globally less tongue body contact than unstressed /a/ (+PI and -CI) for both subjects. Stressed /i/ and /u/, like stressed /a/, are produced with a more posterior tongue contact (+PI). The averaged PI values are: PI= 6.04 (S), PI=4.32 (U). In high vowels, however, also the centrality index tends to increase with stress (+CI). The trend to increase CI in stressed high vowels is consistent, although it does not reach the significance level in all three prosodic contexts (see for EF, the data for /i/ in embedded words (EM), and for /u/ in final words (FN) in Fig.3). Typical palatographic configurations of stressed and unstressed vowels for the two subjects are illustrated in Fig.4.

The context has no systematic effects on the vowel spatial parameters. Specifically, there is no evidence of a consistent trend towards a weakening of the articulatory characteristics of prominence in the stressed vowels of the weakly accented embedded words vs isolated or final words (see, Fig. 3, vowel /a/ for both subjects). There are some effects for the high vowels in the expected direction (reduction of articulatory prominence in embedded words) but they are not shared by the two subjects. For GA posteriority of both stressed and unstressed /u/ decreases in embedded words (p=0.015). For EF, the only effects of context (p = 0.021) occur in stressed /i/, where centrality decreases from CI= 4.7 (isolated words) to CI= 3.9 (embedded words).

3.1.2. Prominence in Consonants

The data on the duration and the maximum EPG contact for the key consonant /t/ are shown for each subject in Fig.5. The analysis indicates that both stress and prosodic context affect the EPG contact and the duration of the consonant. For both subjects CFRONT and C duration are larger in stressed than in unstressed syllables and larger in isolated and final words than in embedded words (p values ranging from 0.001 to 0.000). The average effects of stress are, for CFRONT: 65.9% (S), 57.5% (U), for CDUR: ms 84.63 (S), ms 54.67 (U). The average effects of context are, for CFRONT: 71.5% (IS), 60.97% (FN), 52.70% (EM); for CDUR: ms 91.37 (IS), ms 68.07 (FN), ms 49.51 (EM). Fig 5 clearly shows that stress and context sum up their effects so that consonants in stressed syllables in isolated words have the maximum contact and duration, those in unstressed syllables in embedded words have the minimum. In spite of the wide range of front contact variation, cases of reduced and heavily reduced consonants (from incomplete closure to side contact) were observed only in unstressed syllables of embedded and final words; they amount to 8% for EF and 13% for GA.

3.1.3. Comments on the articulation of stress

The systematic increase in PI (i.e. decrease in anterior side contact) in all three stressed vowels as compared to the unstressed ones, indicates that they are produced with a lower tongue position in front of the constriction. While in stressed /a/ CI decreases (indicating that also the tongue body is lower in stressed than in unstressed /a/), in the production of stressed /i/ and /u/ CI increases, indicating a higher tongue body in the palatal and postpalatal regions. Thus, increase in both parameters (+CI, +PI) indicates that stressed high vowels are articulated with a lower tongue front, and a higher tongue body than the unstressed counterparts. Even if it would be possible to interpret the two sets of data as movements of the tongue proper, i.e. retraction of tongue front associated with elevation of tongue body, the fact that in stressed /a/ the increase in PI is not accompanied by any tongue body elevation, and that in high vowels the increase in PI is larger and more systematic than the increase in CI, is

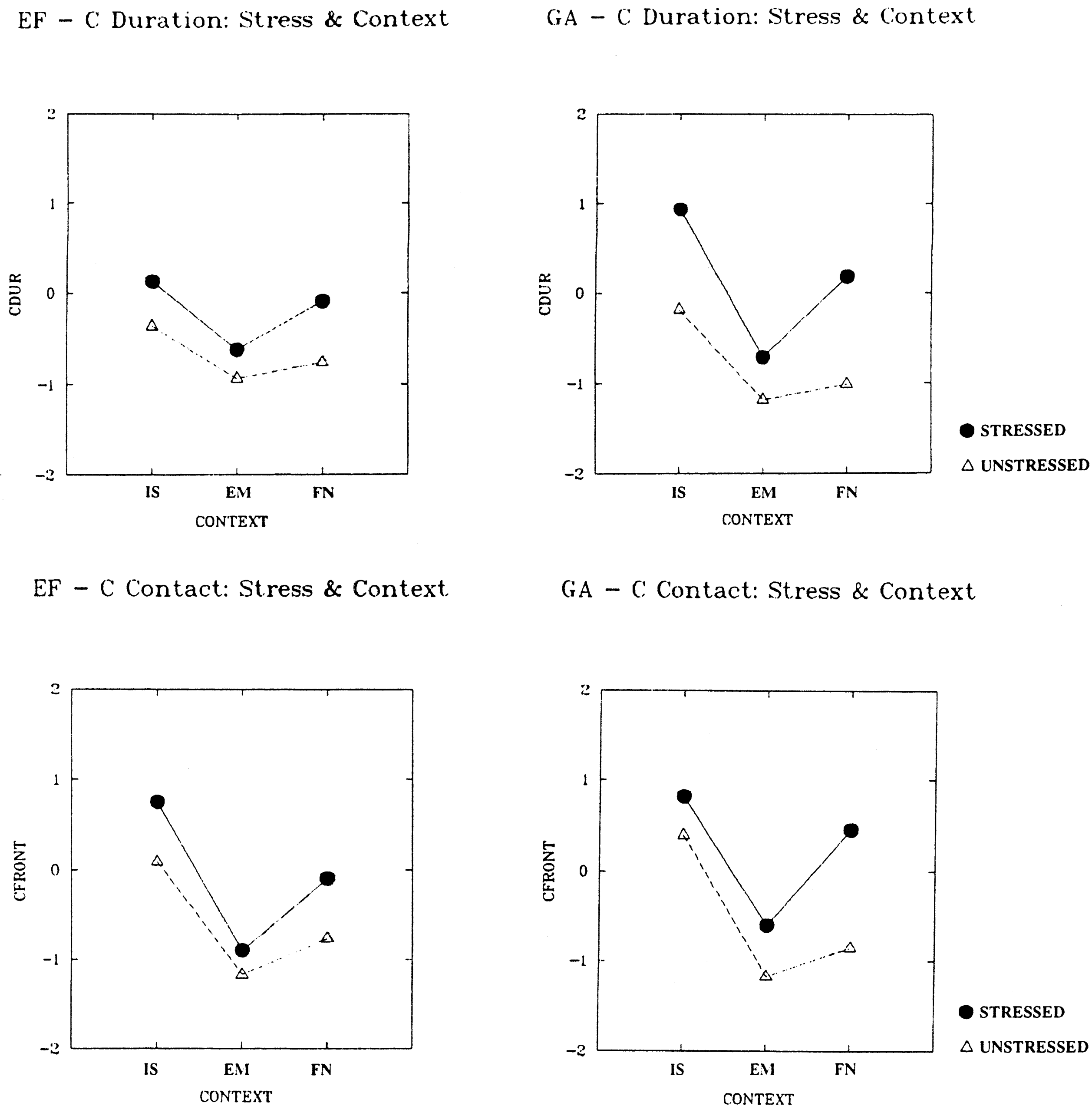

Fig. 5: Mean durations (upper graphs) and mean values of Front Contact (lower graphs) of consonant /t/ as a function of stress and context.

against this hypothesis. Studies on the role of the jaw and of the tongue in the production of prominence (Edwards et al. 1991, de Jong, 1995, among others), and a pilot Movetrack study on vowels /i/ and /a/ in Italian for subject EF (Farnetani and Faber, 1992) lead to an account of the present EPG findings in terms of tongue and jaw articulation. The 1992 data on Italian showed that both stressed /a/ and /i/ were produced with a lower jaw position than the unstressed counterparts. As for the tongue position, it was lower for stressed /a/, but not for stressed /i/. The separation between the jaw and the tongue component in the tongue data, indicated that the observed tongue position was the result of a larger rising movement of tongue articulator in the production of stressed /i/. The two independent movements of the tongue and the jaw in opposite directions in the production of stressed /i/ converge with the present EPG observation. We can thus infer that the increase in posteriority for stressed /a/, /i/, and /u/ be associated with a jaw lowering movement (much larger for /a/ than for /i/ or /u/), while the increase in centrality around the constriction place of high vowels be associated with a more extended tongue rising movements toward the target, and a consequent tighter constriction.
The effects of the prosodic context on the duration of word-medial stressed vowels, which are shorter in embedded vs isolated words for both subjects, indicate that longer segments are more subject to temporal compression than shorter ones, and that also word medial syllables undergo temporal compression in continuous speech. Moreover, the vowel shortening in embedded words vs sentence final words observed for one subject (EF), indicates either a global lengthening of utterance final words, or according to the model of Lindblom and Rapp (1973), a trend towards a durational compression of words as the distance from the utterance boundary increases. The fact that the stressed vowels in embedded words do not systematically undergo reduction in spite of their shorter durations, suggests the activation of compensatory manoeuvres to preserve the spatial quality characterizing prominent vowels.
The consonants parallel the vowels in the durational and spatial changes as a function of stress: syllable prominence is always associated with a tighter, more expanded, and longer consonant closure. Instead, consonants differ from vowels in the effects of context: stressed consonants reduce both their duration and the tongue to palate contact in embedded words. In this condition stressed consonants may be considered less strong than in isolated words, although it must be reminded that in embedded words they are always articulated with complete closure with a contact (average CFRONT= 56%) covering the three first rows, i.e. the entire dentoalveolar region.

3.2. The articulation of boundaries.

In order to test whether the whole CV syllable or only the word initial and final segments are affected by word and utterance boundaries, both C and V were analyzed. Fig.6 and Fig.7 illustrate the global results for the consonant and the vowels, respectively.

3.2.1. Initial boundaries: consonant /t/

The results of paradigmatic comparisons between word initial and word medial consonants (C1 vs C2) indicate that both the closure duration (CDUR) and the C front contact (CFRONT) are larger in word initial position than in word medial position. The data are highly significant for both subjects, and for each prosodic context (p=.04 0.000), as illustrated in Fig.6 (syll. 1 vs syll.2). Moreover the data indicate that the boundary effects on C are significant in both stressed and unstressed syllables.
The interesting finding concerning initial boundaries is that they are signalled also in embedded and in final words: it is reminded that in both conditions the word initial boundaries are phrase internal, hence it appears that initial boundaries are signalled also at the word level. The average differences between C1 and C2, which can be considered to reflect the strength of initial boundaries in the three prosodic contexts are: CDUR = ms 97.84 (IS); ms 26.21 (EM); ms 23.86 (FN); CFRONT = 8.87% (IS); 14.41 (EM); 10.78% (FN). Thus, it seems that in the more constrained condition of phrase-internal boundaries, it is the strengthening of the closure, more than the increase in duration of the silent interval to mark the word initial boundary (see Fig.6). The reason why the very long duration of consonants in absolutely initial position is not paralleled by a comparable increase in contact can be found in the fact that the contact in isolated words is always very high and cannot extend further back once it has reached the maximum extent compatible with the articulation of an alveolar stop consonant (around 82% of the entire front region).

3.2.2. Initial Boundaries: Vowels

The changes in V1 with respect to V2 suggest that also the vowel in syll.1 may contribute to signal

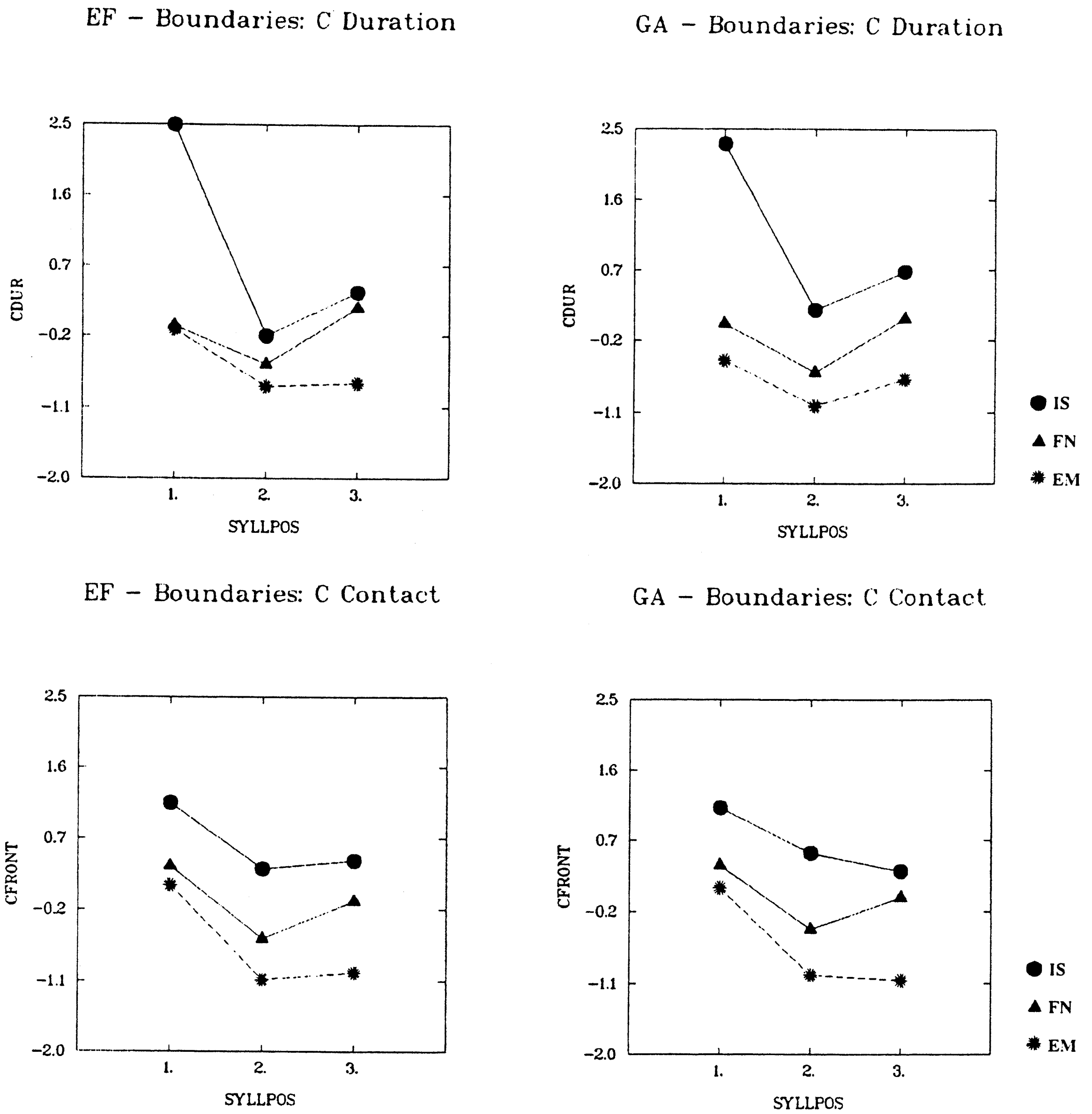

Fig. 6: Mean durations (upper graphs) and mean values of Front Contact (lower graphs) of consonant /t/ as a function of boundaries.

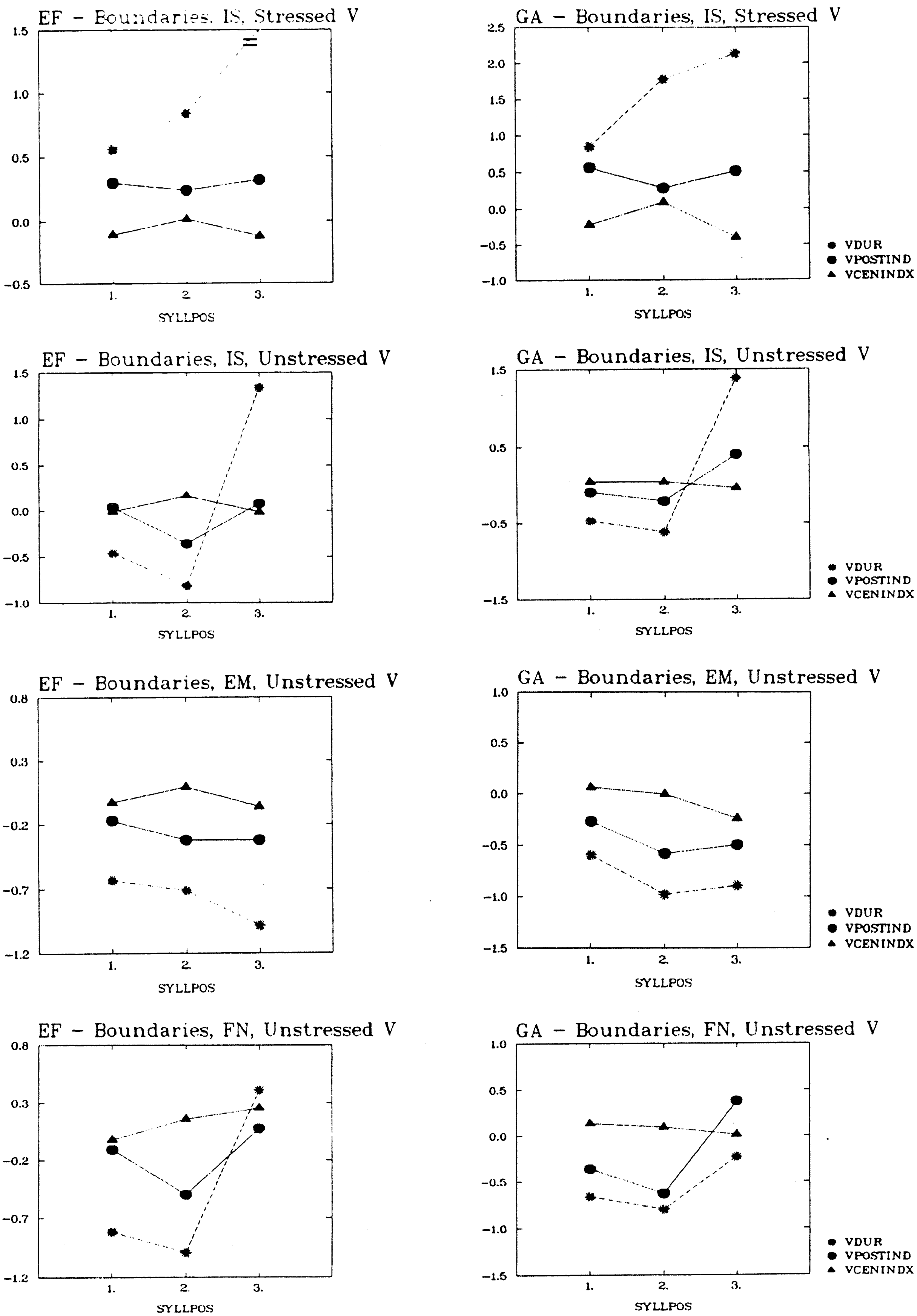

Fig. 7: Average values of vowel duration, PI and CI as a function of boundaries.

initial boundaries. The most systematic spatial change concerns the posteriory index (PI), which is larger in V1 vs V2, and indicates that vowels in syll.1 are produced with an increased opening of the vocal tract.

In isolated words both the stressed and the unstressed vowels tend to have larger PI values syll. 1 than in syll.2, with p=0.000 for EF, p=0.008 for GA (see Fig.7, IS Stressed V; IS Unstressed V, syll.1 vs syll.2, for both subjects). For EF also CI varies significantly and tends to decrease in the first syllable (p=0.013 for stressed Vs; p=.04 0.017 for unstressed Vs). The durational data are peculiar and indicate that in isolated trisyllables stressed vowels in initial position ('CVCVCV words) tend to be shorter than those in medial position (CV'CVCV words) (p=0.010 for EF; p=0.000 for GA), while unstressed vowels are always longer in syll.1 than in syll.2 (p=.04 0.000 for EF; p=.04 0.004 for GA).

In embedded and final words only unstressed vowels vary significantly as a function of boundaries, with PI tending to be higher in V1 than in V2 (p=.04 0.001 for EF; p=.04 0.001 for GA), thus reflecting the trend observed in isolated words. VDUR is higher in V1 vs V2 only occasionally, in embedded words for GA, and in final words for EF (see Fig.7, GA - EM, EF -FN, syll.1 vs syll.2)

The overall data indicate that the vowels of the first syllables (especially unstressed Vs) tend to be more prominent than word medial vowels owing to a larger vocal tract opening often associated with a longer duration. In embedded and final words, these differences (albeit rather small), together with the systematic increase in tongue-front contact of the onset consonant, can contribute to cue to a word-initial boundary in phrase-internal position.

3.2.3. Final boundaries: consonant /t/

The results of comparisons of C2 vs C3 indicate that also the onset consonant of the final syllable contributes to mark final boundaries. Significant changes occur in isolated and sentence final words but not in embedded words. Final boundaries are signalled by significant increase in closure duration in the last syllable : as can be seen in Fig.6, CDUR is longer in syll.3 than in syll.2 (p=0.000 for both subjects). The front contact (CFRONT) varies at a much lesser extent: it tends to be higher in syll.3 than in syll.2 only in sentence final words (p=0.01 for EF; p=0.047 for GA)

3.2.4. Final boundaries: Vowels

The overall results indicate that the word final boundaries are signalled in isolated and in sentence final words, as expected, but not in embedded words. (see Fig.7, Syll.2 vs syll.3 : IS vs EM vs FN for both subjects).

Final boundaries are signalled by a significant increase in duration of final vowels. As illustrated in Fig.7, V3 is longer than V2 for both subjects (p=0.000); the average durational values are ms.56.4 (V2) and ms. 111.54 (V3). The increment in duration of the final vowel is greater in unstressed than in stressed syllables (see, Fig.7, stressed V vs unstressed V in Isolated Words).

Also vowel posteriority increases significantly in syll.3 vs syll.2 in isolated and final words; the average PI values are PI=4.34 for V2, PI= 5.77 for V3. The increment in PI is more consistent in unstressed than in stressed vowels. For GA it is significant in both stress conditions (p=.04 0.018 for Str.V; p=0.000 for Uns.V;), for EF only for unstressed vowels (p=.04 0.003). As can be seen in Fig.7, the changes in PI and in VDUR in final vowels are similar to those induced by initial boundaries, but are much larger for both parameters. It has to be noted that in embedded words, where the boundary is not realized, the duration of the word final vowel tends to be even shorter than the word medial vowel (see Fig.7, EM, Subj. EF).

3.2.5. Comparing the articulation of stress and boundaries

To a certain extent, the articulation of boundaries resembles the articulation stress: duration and front contact increase in consonants, posteriority increases in vowels, accompanied by a relevant increase in duration in absolutely final vowels. These combined effects on C and V make initial and final unstressed syllables more prominent than word-medial unstressed ones, while a stressed syllable at the word edges further enhances the prominence contrast at the word level. But there are important differences between the effects of stress and those of boundaries, first, quantitative differences: in vowels, the spatiotemporal changes induced by boundaries are smaller than those induced by stress (compare Fig. 2 and Fig.3 with Fig.7), and, second, stress induces also an increase in centrality in high vowels, which indicates that the tongue body has a higher position and a tighter constriction. Boundaries do not induce such effects, on the contrary, when CI changes, it tends to decrease at the boundaries. Thus, the articulation of boundaries comports mainly a larger opening of vowels and a

tighter closing of consonants, whose outcome is an enhancement of the V/C contrast. Instead, the articulation of stress, comports not only an expansion of vowel opening and consonant closing movements, but also an increase in the target-directed movement for each vowel, hence, an enhancement of the vowel specific distinctive properties and a consequent increase in the phonetic distance among the different vowels.
As for vowel duration, while final boundaries induce an increase in duration of both stressed and unstresssed final vowels, with stronger effects on unstressed vowels (as can be seen in Fig.7 for isolated words), initial boundaries induce a lengthening of unstressed vowels only (see Fig.7 all contexts). The stressed vowels of syll.1 ('CVCVCV words) are never longer than those in syll.2 (CV'CVCV words), on the contrary, they are significantly shorter in isolated words (see Fig.7, IS). These data on stressed vowel duration agree with a previous study on duration as a function of word size and stress position (Farnetani and Kori, 1986). In that study the data indicated that stressed vowels are longer in 'CVCV words than in 'CVCVCV but not in CV'CV vs CVCV'CV words, which showed that the durational differences were not related to word size but rather to the number of unstressed syllables following the stressed one, in other words the stressed Vs shorten as the word extends rightwards. Thus, in the present study, the shorter duration of stressed vowels in the initial syllables of 'CVCVCV words vs the medial syllables of CV'CVCV words are to be related to the greater distance of the stressed syllable from the end of the word, and can be accounted for by rhythmical constraints.

4. General discussion and conclusion

The first purpose of the present research was to know to what extent the prosodic features of lexical stress and boundaries influence the articulation of a segment. The data indicate that the articulatory spatial changes are systematic, quantitatively relevant and that in most cases they are not mere consequences of durational changes, but rather active articulatory modifications aiming at the production of various degree of articulatory prominence. As seen before, in the production of stress and boundaries both the consonant and the vowel of the CV syllable play an important role. As for the effects of the prosodic context, the data have shown that lexical prominence is not weakened when duration is compressed in words embedded within a prosodic phrase, and that in phrase internal position the word initial boundaries are preserved, although weakened, while word final boundaries are totally cancelled. Thus it appears that the syllabic prominence induced by lexical stress and, to a lesser extent, the prominence resulting from the articulation of initial boundaries are maintained also at the word level.
The second purpose of this study was confront the present data with the models proposed for the production of prominence and boundaries. The inferences that can be drawn from the analysis of the spatial configuration of C and V are that lexical prominence, initial boundaries, and final boundaries, though similar in their articulation under certain respects, differs substantially among each other: it has been shown that the articulation of stressed syllables is characterized by a strengthening of C, an expansion of vowel opening and an enhancement of vowel specific articulatory features, and this would agree with the hypothesis that a stressed syllable is a hyperarticulated syllable. This cannot be said, at least for high vowels, for the articulation of boundaries: both initial and final boundaries do induce induce a strengthening in the consonant but only an expansion of opening in the vowels. Moreover there are substantial differences between initial and final boundaries: initial boundaries are marked articulatorily also when duration does not increase, while final boundaries are marked more strongly by duration than by an increase in articulatory prominence.
In order to gain some more knowledge on the control strategies underlying the production of stress and boundaries, multiple regressions have been carried out for vowel /a/. Even if tongue movements cannot be recorded with EPG, they can be indirectly inferred from the traces left by the tongue-to-palate contact over time. In this analysis the number of free electrodes in vowel /a/, measured in the first temporal frame of minimum contact was the dependent variable of a number of stepwise regressions with the amount and the duration of the C-to-V displacement, the mean displacement rate and the maximum (or peak) displacement rate were the independent variables. The purpose was to infer how prominence, initial boundaries and final boundaries are produced in terms of C-to-V gestures. The characteristics of C-to-V gestures for stress were analyzed in word medial syllables, those for initial and final boundaries were analyzed by confronting word medial syllables with word initial and word final syllables, respectively. The preliminary results indicate that the very low contact characterizing stressed /a/ is in general accounted for by an increased duration and displacement of the C-V movements. But in

embedded words stress is associated with the duration and the peak rate of the C-V movement. Initial boundaries production (unstressed syllables) is associated with the duration and the rate of the movements for GA and to the their peak rate for EF, while the articulation of final boundaries (unstressed syllables) seems to be accounted for only by the duration of the C-V movements for GA and by displacement and duration for EF. This preliminary data, in which there is a good agreement between the two subjects, suggest that hyperarticulation characterizing stressed vowels may be simply the outcome of a sufficient time allowed to the articulators to achieve the target; however, in embedded words where segments undergo temporal compression the stressed vowel is still hyperarticulated, and in this case the target is achieved by increasing the rate of the gestures. The two different articulatory strategies inferred from the present data are compatible with the revised H&H model (Moon and Lindblom, 1994) were *duration* and *input force* are two of the predictors of vowel undershoot and reflects the strategies available to speakers under different circumstances. Hence, the sonority expansion hypothesis (Beckman et al.1992) according to which prominence is brought about by decreasing the temporal overlap between adjacent gestures, can account only in part for the present data. As for the articulation of the vowel in the first syllable of the word, where the spatial changes are larger than the temporal ones, the articulatory movements must be faster in order to attain a greater opening of the vocal tract in the production of an unstressed vowel without interfering with the rhythm imposed by the stressed/unstressed contrast. Final boundaries seem to be articulated by slowing down the speech tempo (which could account for the very long vowel duration) and/or by expansion of the jaw movements (which could account for a an increase in contact for the consonants and an increase in opening for the vowels). Further analysis is now in progress for high vowels.

References

Beckman, M.E., Edwards, J.R., and Fletcher, J. (1992). Prosodic structure and tempo in a sonority model of articulatory dynamics. In G. Docherty and D.R. Ladd (eds), *Papers in Laboratory Phonology II, Segment, Gesture, and Tone*. CUP, pp. 68-86.

Browman, C.P., and Goldstein, L. (1989). Tiers in articulatory phonology, with some implications for casual speech. In J. Kingston and M.E. Beckman (eds), *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech.* CUP, pp. 341-376.

de Jong, K.J. (1995). The supraglottal articulation of prominence in English. Linguistic stress as localized hyperarticulation. JASA 97, 491-504.

Edwards, J.R., Beckman, M.E., and Fletcher, J. (1991). The articulatory kinematics of final lengthening. JASA 89, 369-382.

Farnetani, E. (1989). Acoustic correlates of linguistic boundaries in Italian: A study on duration and fundamental frequency. Proceed. of European Conference on Speech Communication and Technology. Eurospeech 89, Vol.2, 332-335.

Farnetani, E., and Kori, S.(1983). Interaction of syntactic structure and rhytmical constraints in the realization of word prosody. Quaderni CSRF 2, 288-318.

Farnetani, E. and Kori S. (1986). Effects of syllable and word structure on segmental durations in spoken Italian. Speech Communication 5, 17-34.

Farnetani, E. and Kori, S. (1990). Rhythmic structure in Italian noun-phrases: A study on vowel duration. Phonetica 47, 50-65.

Farnetani E. and Faber A. (1992). Tongue-jaw coordination in vowel production: Isolated words vs connected speech. Speech Communication 11, 401-410.

Farnetani, E. and Vayra, M. (1996). The role of prosody in the shaping of articulation in Italian CV syllables. Proceed. 1st ESCA Tutorial and Research Workshop on Speech Production Modeling. Autrans, 9-12.

Fougeron, C. and Keating, P. (1996). The influence of prosodic position on velic and lingual articulation in French: Evidence from EPG and airflow data. Proceed. 1st ESCA Tutorial and Research Workshop on Speech Production Modeling, Autrans, 93-96.

Keating, P. (1995) Effects of prosodic position on /t,d/ tongue/palate contact. Proceed. ICPhS 95, Stockholm, V3, 432-435

Keating, P. (in press). Word-level phonetic effects on English consonant articulation . Paper presented at the Conference on the Phonological Word. ZAS, Berlin, October 1997.

Kohler, K.J. (1990). Segmental reduction in connected speech in German: phonological facts and phonetic explanations. In W.J. Hardcastle and A. Marchal (eds), *Speech production and speech modeling*. Dordrecht: Kluwer Academic Publishers, pp. 69-92.

Lea, W.A. Prosodic aids to speech recognition (1980). In W.A.Lea (ed) *Trends in Speech recognition*, Prentice Hall Inc. New Jersey, pp. 166-205.

Lindblom, B. (1990). Explaining phonetic variation: a sketch of the H&H theory. In W.J. Hardcastle and A. Marchal (eds), *Speech Production and Speech Modelling*, Kluwer Academic Publishers, pp. 403-439.

Lindblom, B. and Repp, K. (1973). Some temporal regularities of spoken Swedish. Papers from the Institute of Linguistics, Stockholm University, 21, 1973, Monograph

Macchi, M. (1985).Segmental and suprasegmental features and lip and jaw articulators. Ph.D. dissertation, New York University.

Moon, S.J. and Lindblom, B. (1994). Interaction between duration, context, and speaking style in English stressed vowels. JASA 96, 40-55.

Perkell, J.S. (1997). Articulatory Processes. In W.J.Hardcastle and J.Laver (eds), *The Handbook of Phonetic Sciences,* Blackwell Publishers, pp.333-370.