Phonetica
Editor: K. Kohler, Kiel

Separatum
Publisher: S. Karger AG, Basel
Printed in Switzerland

811

Original Paper

Phonetica 1992; 49: 48–60

Mario Vayra[a]
Carol A. Fowler[b, c]

[a] Scuola Normale Superiore, Pisa, Italy;
[b] Haskins Laboratories, New Haven, Conn., USA;
[c] Dartmouth College, Hanover, N. H., USA

Declination of Supralaryngeal Gestures in Spoken Italian

Abstract

Two experiments investigate a weakening of supralaryngeal gestures in an utterance, analogous in some ways to declination of fundamental frequency and amplitude. In one experiment, acoustic measures revealed progressive centralization of stressed /i/, /a/ and /u/ left to right in trisyllabic utterances read by Tuscan subjects. A second experiment, using speakers of a different (Northern) variety of Standard Italian, found reduction in jaw opening for stressed /a/ left to right, but generally failed to replicate a centralization of /i/. This experiment further suggested that the progressive weakening of supralaryngeal gestures is largely a phrase level rather than a word level phenomenon. Both experiments found a different, V-shaped, pattern of opening to be generally characteristic of unstressed syllables.

In an investigation of coarticulation in spoken Italian [Vayra and Fowler, 1987] we obtained an unexpected finding. In that study, 3 speakers of the Florentine variety of Standard Italian had produced bi- and trisyllabic nonsense words in isolation. Initial and final vowels were /i/, /a/ and /u/; the medial vowel of the trisyllables was stressed or unstressed /a/. Acoustic formant measures revealed that a major source of variation in stressed /a/ was due to its position in a bi- or trisyllable. In particular, aside from any coarticulatory effects of neighboring vowels, all 3 speakers exhibited a considerably higher F_1 of stressed /a/ in initial than final position in bisyllables and a monotonic decrease in F_1 of stressed /a/ in trisyllables. The pattern on unstressed syllables was a V shape with the medial syllable least open.

Received:
May 18, 1990
Accepted:
October 29, 1991

Prof. Carol A. Fowler, Dartmouth College
Department of Psychology,
Hanover, NH 03755 (USA)
Dr. Mario Vayra, Scuola Normale Superiore
Piazza dei Cavalieri, 7, I–56100 Pisa (Italy)

Table 1. Mean first formant frequencies of /a/ in Hertz and standard deviations (SD) for 3 speakers (S, F, N) as a function of position in a trisyllable

Speakers		Stressed			Unstressed		
		I	M	F	I	M	F
S	mean	1,178	1,112	989	1,144	803	850
	SD	5.2	25.4	36.9	52.78	23.2	3.06
F	mean	1,052	986	890	972	794	859
	SD	13.2	40.1	26.3	22.2	11.7	30.8
N	mean	877	800	725	794	701	705
	SD	37.8	19.4	27.1	13.8	19.1	20.7

I = Initial; M = medial; F = final.

Table 1 shows the findings on trisyllables from the 3 speakers (S, F, N). We proposed tentatively that the pattern on stressed syllables might reflect a process analogous in some ways to declination of fundamental frequency (f_0) in which effects of stress decrease over the course of an utterance.

In declination of f_0, particularly in accented syllables [e.g., Breckenridge, 1977; Cohen et al., 1982], f_0 decreases from early to late in a phrase. Several accounts of declination have appeared in the literature [e.g., Breckenridge, 1977; Cooper and Sorensen, 1981; Gelfer, 1987; Maeda, 1976] and it is likely that several factors contribute to its occurrence. A fundamental one may be regulation of subglottal pressure [P_s; e.g., Gelfer, 1987; Gelfer et al., 1987]. A major influence on f_0, besides tension in the intrinsic muscles of the larynx, is P_s, and one source of influence on P_s is the elastic recoil force of the lungs. As the lungs deflate, the elastic recoil force decreases and so, other things equal, does P_s, and so, other things equal, does f_0. [For another account also implicating P_s, see 't Hart et al., 1990.] Other things are not equal during speech, and so f_0 does not invariably track the decline in lung volume. However, even so, declination occurs commonly across utterances and languages, although it may not be equally common in all speaking styles [Lieberman et al., 1985; Umeda, 1982]. Compatible with an idea that declination of f_0 has its origins in the respiratory system are findings that amplitude falls along with f_0 [Breckenridge, 1977].

For its part, a monotonic decrease in F_1 of stressed /a/ early to late in an utterance suggests a centralization of stressed /a/ in the height dimension. Is it possible that speech exhibits a *general* tendency to weaken early to late in an utterance? If so, declination of f_0 is not an independent characteristic of a speech utterance as it has been treated, but rather, it is an index of a general weakening tendency.

Further evidence that declination may be very general in speech production is provided by Bell-Berti and Krakow [1991], who measured velum height during consonants associated with high velar position (/s/, /t/). They varied position of these consonants in a sentence and found progressive early to late e reductions of velum height during the consonants. The reduction was considerably more marked for stressed as compared to unstressed syllables (in which velum height was already rather low for /s/ and /t/ as compared to their height in stressed syllables).

In Italian, a declination effect seems to be present in Bertinetto's [1981, pp. 65–80] data on intensity and f_0 peak values in stressed and unstressed vowels in initial, medial and final position of isolated trisyllables [see his table

5]. In stressed vowels, declination of intensity showed up clearly only when intensity was calculated as the area under the intensity curve [compare his tables 4 and 6].

Additional evidence of intensity declination (with intensity now measured as peak values of the intensity curve) is present in an earlier study by Fava and Magno Caldognetto [1976] on Italian isolated bisyllables of the types /'CVCV/ versus /CV'CV/ and /'CVCCV/ versus /CVC'CV/. In that paper, the role of duration and intensity in distinguishing stressed and unstressed vowels was investigated. Results showed a clear position effect for intensity, with greater intensity always associated with the initial syllables of the bisyllables, whether or not that syllable was stressed.

A final measure that we will review here is duration. It is not entirely clear how declination in duration – either articulatory of acoustic – should manifest itself. One possibility is that declination in production of articulatory gestures will reflect progressive reduction of gestures and, therefore, shortening of their articulatory extents and acoustic consequences. Alternatively, however, declination may be associated with slowing and, therefore, lengthening of gestures and their acoustic products (as in final lengthening, utterance-finally). In Italian, studies of acoustic durations generally show a shortening of final, as compared to nonfinal stressed vowels [see, for example, Farnetani and Kori, 1982; Marotta, 1985]. This outcome is compatible with the traditional phonological description of Italian's stressed vowels' length in word-final position [for example, Vincent, 1988; Basbøll, 1989]. Notice, however, that the evidence does not indicate either progressive durational reduction or lengthening; rather, in trisyllables, there may be an overall tendency for medial syllables to be longest [Vayra et al., 1987].

There is also a position effect on the perceptual side when stress judgements are made on synthesized bisyllabic words, as reported by Bertinetto [1980]. He synthesized a total of 64 different bisyllabic stimuli corresponding to the phonological sequence /papa/, which in Italian can give rise to two different words, according to stress position (/'papa/ 'Pope' versus /pa'pa/ 'father'). Notice that, as the author remarks, 'no effort was made to preserve just those characteristics of speech which are specifically attached in normal speech to different syllabic positions' [p. 386]. That is, the contrasts among the 64 synthesized /papa/ stimuli of the experiment were only realized through the systematic assignment of four different values of duration, f_0 and intensity to both syllables of each sequence. In the experiment, more stimuli were perceived as stressed on the second syllable than as stressed on the first; moreover, for any parametric value (duration, f_0, intensity) assigned to a syllable, there were more stress judgements in favor of final as compared to initial syllables.

It seems to us that, under a hypothesis of declination as a general articulatory process involving a decrement in the values of different acoustical parameters, this imbalance in stress judgements may indicate that listeners 'expect' global declination in speech. In the same way, Pierrehumbert [1979; see also Silverman, 1987; Terken, 1991] has shown that listeners 'expect' f_0 declination to occur. They judge two intonational peaks of a sentence as equal only if the later one is sufficiently lower than the earlier one. Thus, the subjects of Bertinetto's experiment parse the global prosodic contour of the stimuli in at least two components – declination and stress – by factoring out information on stress from the expected weakening of the articulatory system [see Fowler, 1986; Fowler and Smith, 1986, on the principles underlying a style of perception by 'vector analysis'].

In the present study, we begin to test the idea that, in Italian speech, declination extends beyond effects on f_0 and perhaps amplitude as well. We begin with further analysis of the speech of the 3 speakers examined in Vayra and Fowler [1987]. If speech does exhibit general weakening, then evidence of it may be apparent in aspects of their utterances not yet examined.

These speakers produced bi- and trisyllables beginning and ending with stressed or unstressed /i/, /a/ and /u/. The medial syllable of the trisyllables was stressed or unstressed /a/. If speech weakens over an utterance, then /i/ and /u/ should exhibit evidence of progressive centralization. If the effect is weakening of stress, then any centralization we see may be limited to stressed syllables. Whereas stressed /a/ showed a decrease in openness (a lowering of F_1), stressed /i/ and /u/ should show progressive opening if they centralize. In addition, vowels may show progressive centralization along the front-back dimension. Accordingly, F_2 of /i/ may lower while F_2 of /u/ increases.

Experiment 1

In this experiment, described in more detail in Vayra and Fowler [1987], speakers were 2 female and 1 male speaker of the Florentine variety of Standard Italian. They produced bisyllabic and trisyllabic nonsense words in isolation. In each nonsense word, one syllable was stressed; across words, stress occurred equally often in each syllable of the bi-and trisyllables. Initial and final vowels were /i/, /a/ and /u/; they appeared in all combinations across bi- and trisyllables. The medial syllable was always /a/. Accordingly, there were 18 different bisyllables and 27 different trisyllables. Each speaker produced each nonsense word three times.

Speech was filtered at 5 kHz and digitized at a 10-kHz sampling rate. Vowel formant values were estimated using linear predictive analysis of the ILS system at Haskins Laboratories [ILS Command Reference Guide V. 6.0, 1986]. A 20-ms analysis window was moved in 5-ms steps over the trisyllables. The frequency spectrum for each analysis frame was obtained by performing an FFT on the autoregressive coefficients calculated from the regression coefficients for each frame. We took measures at the vowel midpoints.

Results and Discussion

Tables 2 and 3 provide F_1 and F_2 measures for stressed and unstressed initial and final syllables averaged over the bi- and trisyllables as produced by the 3 speakers. On average, every comparison conforms to a prediction that vowels centralize early to late in an utterance. For both /i/ and /u/, F_1 increases, suggesting a progressive opening over time of these close vowels. For /i/, this is significant for 2 of the 3 talkers [S: $F(1, 82) = 107.04$, $p < 0.001$; F: $F(1, 82) = 114.9$, $p < 0.001$]. It is significant for the same 2 talkers for the vowel /u/ [S: $F(1, 82) = 4.88$, $p = 0.03$; F: $F(1, 82) = 15.3$, $p < 0.001$]. For subject S producing /u/, there is a three-way interaction with stress and word length because the effect fails to hold for stressed syllables in trisyllables.

As for F_2 of /i/, it decreases early to late in the bi- and trisyllables, suggesting that this front vowel shifts back. This effect is significant for all 3 talkers [S: $F(1, 82) = 265.46$, $p < 0.001$; F: $F(1, 82) = 34.3$, $p < 0.001$; N: $F(1, 82) = 71.29$, $p < 0.001$]. As for F_2 of /u/, it increases, indicating either fronting or decreased rounding. This difference is significant for 2 talkers [F: $F(1, 82) = 35.87$, $p < 0.001$; N: $F(1, 82) = 12.1$, $p = 0.001$]. The 3rd talker shows a three-way interaction with stress and word length; the increase in F_2 holds for unstressed syllables in unstressed bi- and trisyllables and in stressed trisyllables.

In sum, we found considerable evidence of centralization, but no indication that it is restricted to stressed syllables. Accordingly, the F_1 minimum that all talkers showed for un-

Table 2. Mean first formant frequencies of /i/ and /u/ in Hertz and standard deviations (SD) for 3 speakers (S, F, N) as a function of position in a trisyllable

Speakers		Stressed		Unstressed	
		I	F	I	F
First formant /i/					
S	mean	248	298	243	290
	SD	19.0	21.8	13.4	21.9
F	mean	253	270	310	330
	SD	22.8	29.7	27.0	26.4
N	mean	255	253	262	260
	SD	28.1	37.4	33.6	34.1
First formant /u/					
S	mean	296	296	305	340
	SD	21.8	43.3	22.5	33.8
F	mean	307	343	336	363
	SD	20.1	35.3	46.5	24.1
N	mean	250	260	274	262
	SD	30.5	28.4	33.7	30.7

I = Initial; F = final.

Table 3. Mean second formant frequencies of /i/ and /u/ in Hertz and standard deviations (SD) for 3 speakers (S, F, N) as a function of position in a trisyllable

Speakers		Stressed		Unstressed	
		I	F	I	F
Second formant /i/					
S	mean	2,938	2,739	2,876	2,618
	SD	53.3	97.1	55.2	57.6
F	mean	2,563	2,477	2,485	2,367
	SD	56.2	69.6	99.5	87.3
N	mean	2,240	2,125	2,222	2,098
	SD	62.3	56.7	41.7	81.8
Second formant /u/					
S	mean	721	711	713	787
	SD	70.5	55.6	89.3	71.2
F	mean	727	800	711	847
	SD	58.6	55.6	89.3	71.2
N	mean	554	604	563	530
	SD	61.4	69.3	78.3	54.2

I = Initial; F = final.

stressed medial /a/ remains unexplained in terms of a general tendency for vowels to centralize. We explore the finding further in experiment 2.

Experiment 2

The next experiment was designed to further pursue our study of supralaryngeal declination. The first experiment does not clarify whether the centralization effects we observed are word level effects or, like declination of f_0, occur over phrasal domains. If they are word level effects, then they should be as evident when nonsense words are produced in a carrier phrase as when they are produced in isolation. If they are phrase level effects, they should be considerably reduced or even absent when the words are produced in a carrier phrase. A second issue concerns the coupling of centralization to declination ostensibly having an origin in the respiratory system. Accordingly, we asked whether these effects are accompanied by declination of f_0 and amplitude, and, if they are when words are produced in isolation, whether the two sets of effect separate out when words are produced in a carrier phrase. The present experiment also added a kinematic measurement to the spectral measures of experiment 1. This allowed us to ask whether declination of F_1 of stressed /a/ is accompanied by kinematic evidence of lesser opening for that vowel.

Method

Subjects were 2 Northern speakers of Standard Italian; 1 was the first author of the paper. The speakers produced trisyllables /bibabi/ and /bababa/ with primary stress on initial, medial or final syllables. Talkers produced 20 tokens of each isolated trisyllable and 20 tokens of trisyllables produced in the carrier phrase: *dico la parola ___ con cura.*

In addition to the acoustic measures taken as previously described in experiment 1, we obtained acoustic durational measures and measures of jaw opening for /a/. Measures of jaw vertical displacement were gathered with a Selspot optical system modified at Haskins Laboratories. Miniature light-emitting diodes were taped midsagittally onto the forehead, nose and point of the chin. Modulated light from the diodes was monitored by a camera equipped with a Schottky planar diode located in its focal plane. The output of the photodiode was input to associated in-house electronics [Gulisano, 1982], translating the signal into X and Y coordinates. In order to eliminate effects of extraneous movement, the outputs of the reference light-emitting diodes (forehead, nose) were subtracted from that of the chin [see Kay et al., 1985, for a more complete description of the signal processing and analysis sequence]. Jaw measurements were taken at the point of jaw maximum opening; measurements are in millimeters.

Simultaneous acoustic recording was made. Acoustic waveforms were displayed and vowel duration measured via a waveform editing program. In absence of a clear burst marking the release of closure for /b/, the duration of voicing for the vowel was measured as the interval from the first acoustic evidence of high frequency components in the periodic wave to the first acoustic evidence of closure (defined as the point where high frequency components disappeared and the amplitude of the waveform dropped markedly). Formant trajectories for the vowels, as well as f_0 and amplitude information were obtained using ILS procedures as described in the previous experiment. Measurements of F_1, F_2, f_0 and amplitude were taken from the vowels' temporal midpoint.

Results[1]

/a/: No Carrier

Table 4 shows the data for isolated productions of the trisyllables including only the vowel /a/. Data are presented separately for

[1] In the 'Results' we report the results of multiple analyses of variance performed on the various dependent variables of the experiment. Before doing that, however, we briefly present the results of four preliminary analyses prompted by a reviewer's suggestion. The reviewer pointed out that a preferable analysis to the ANOVAs we report in the text is a Multivariate

the stressed and unstressed syllables and for the 2 talkers. Both stressed and unstressed syllables exhibit a monotonic decrease in f_0 from early to late. The effect was strong and significant for both talkers[2] [CA: stressed syllables, F(2, 57) = 77.01; unstressed, F(2, 117) = 70.6; MV: stressed, F(2, 57) = 59.81, unstressed, F(2, 117) = 12.3]. Neither talker showed a lower f_0 for unstressed syllables. However, according to Bertinetto [1978, 1980, 1981], f_0 is not a reliable correlate of stress in Italian either in production or perception. [See Ferrero, 1972, for a different view.]

These utterances also exhibited declination of amplitude early to late [CA: stressed, F(2, 57) = 79.52, unstressed, F(2, 117) = 119.8; MV: F(2, 57) = 21.81, unstressed F(2, 117) = 5.16, p = 0.007]. Amplitudes of unstressed syllables were consistently lower than those for stressed syllables for both talkers. This is consistent with comparative work on vowel reduction by den Os [e.g. 1985, 1988] and Koopmans-van Beinum [1980, 1983].

Turning to measures of vowel openness, both talkers showed the same qualitative patterns of F_1 as the 3 talkers of experiment 1. F_1 decreased monotonically early to late in stressed syllables, but the decline was significant only for CA [F(2, 57) = 6.24, p. = 0.004]. Differences in F_1 of unstressed /a/ across positions were significant for both talkers; as in the findings of Vayra and Fowler [1987], the F_1 minimun was in medial, rather than final, unstressed syllables [CA: F(2, 117) = 84.94, p = 0.004; MV: F(2, 117) = 80.19].

Measures of jaw opening replicated findings on F_1. Both talkers showed a monotonic decrease in opening early to late in stressed syllables and a V-shaped pattern for unstressed syllables [CA: stressed, F(2, 56) = 9.21, unstressed: F(2, 115) = 48.07; MV: stressed, F(2, 57) = 72.11, unstressed, F(2, 117) = 94.18]. Most likely, the acoustic decrease in openness of the vowel exhibited by stressed /a/ in the present experiment and those in experiment 1 reflects a progressive decrease in jaw opening for the vowel when it is stressed. The already considerably centralized unstressed vowels do not show additional centralization. They do,

---

Analysis of Variance (MANOVA). This test can be used when the data include multiple dependent as well as independent variables, whereas ANOVA is restricted to just a single dependent variable. The MANOVA allows discovery whether independent variables have effects jointly on several dependent measures. Because we expect that this analysis will be unfamiliar to most readers and because we had to go on anyway to look at the dependent measures separately, we place the result of the MANOVAs here rather than in the body of the text. We performed four MANOVAs, two per stressed vowel per talker. Dependent variables were f_0, amplitude, F_1 and F_2. Independent variables were Carrier, Stress and Position of the vowel in the word. In addition, we used vowel duration as a covariate in the analysis. The same reviewer had suggested that any early to late spectral reduction we observed could be an artifact of a tendency to speed up over the course of word production. Although our durational measures do not bear out this idea, in any case, we used duration as a covariate to allow the test to discover whether there are effects of position of the vowel in the word when effects of duration on the dependent measures are partialed out.

Briefly, in all four MANOVAs, every main effect and interaction was significant, with all p values but one (p = 0.004) falling below the 0.001 level of significance. This indicates that the various dependent measures are affected jointly by the independent variables. Moreover, the effects of those independent variables are not mediated by effects of duration. However, the pattern of results is also highly complex with even the three-way interactions highly significant. Further, a significant multivariate F does not mean necessarily that all dependent measures are affected significantly by the independent variables. Accordingly, in the text we report separate Position x Stress ANOVAs on the two context conditions (No carrier, Carrier) and on the individual dependent measures of the experiment.

[2] Global analyses were initally performed with factors stress and carrier phrase. However, the effect of position in the word quite generally interacted with those variables. Therefore separate analyses were performed on stressed and unstressed syllables and on utterances produced in isolation and in a carrier phrase.

Table 4. Means and standard deviations (SD) for the various dependent measures of experiment 2, vowel /a/ in the initial, medial and final positions of a trisyllable

		Stressed			Unstressed		
		I	M	F	I	M	F
No carrier CA							
f_0, Hz	mean	215	182	177	227	207	175
	SD	12.5	9.0	9.2	11.3	27.3	16.0
RMS, dB	mean	55.8	51.4	50.1	54.6	50.1	44.1
	SD	0.8	1.92	1.5	1.6	3.6	3.4
F_1, Hz	mean	1,052	1,027	957	719	677	868
	SD	94.6	83.2	8.53	53.0	77.3	72.7
Jaw, mm	mean	10.1	8.69	8.38	4.41	3.52	5.93
	SD	1.3	1.2	1.5	1.2	1.0	0.9
Duration, ms	mean	170.9	203.9	144.3	95.2	81.8	99.7
	SD	9.0	12.3	17.1	6.1	10.3	16.0
MV							
f_0, Hz	mean	121	106	95	123	113	111
	SD	10.6	6.4	4.5	6.4	10.6	15.2
RMS, dB	mean	53.7	52.7	47.9	49.5	47.2	45.1
	SD	2.5	2.2	3.7	2.8	3.5	5.4
F_1, Hz	mean	728	720	712	615	547	627
	SD	30.6	21.9	36.2	18.9	27.89	30.8
Jaw, mm	mean	7.79	5.9	3.76	5.67	2.01	2.06
	SD	1.1	1.3	0.8	0.8	0.6	0.6
Duration, ms	mean	139	183	109	73	69	102
	SD	11.3	10.9	22.4	10.3	10.2	21.0
Carrier CA							
f_0, Hz	mean	178	169	175	190	189	231
	SD	8.3	6.1	9.0	7.2	9	17.5
RMS, dB	mean	53.0	51.1	50.6	52.5	51.5	52.7
	SD	1.9	1.7	2.3	1.8	1.7	2.1
F_1, Hz	mean	1,033	1,024	1,051	705	692	902
	SD	71.8	46.3	99.7	59.4	56.9	65.3
Jaw, mm	mean	9.18	9.32	8.89	2.74	3.64	5.8
	SD	1.0	1.2	1.2	0.8	1.0	1.2
Duration, ms	mean	184	202	191	84	85	110
	SD	9.2	11.1	17.4	7.5	11.2	15.9
MV							
f_0, Hz	mean	133	124	124	102	128	136
	SD	1.3	8.9	6.9	4.9	20.5	8.4
RMS, dB	mean	54.3	53.1	52.4	47.5	47.3	50.3
	SD	1.0	1.3	1.1	2.3	7.4	2.5
F_1, Hz	mean	704	736	725	557	529	532
	SD	59	43.6	44.9	56.5	45.7	31.0
Jaw, mm	mean	6.01	4.32	3.12	2.49	1.36	1.12
	SD	0.7	0.9	0.7	0.6	0.4	0.4
Duration, ms	mean	150	167	205	72	72	81
	SD	6.3	12.2	63.3	11.2	10.2	10.1

I = Initial; M = medial; F = final.

however, show a consistent pattern in which the medial syllable is least open.

Finally, we looked at the measures of acoustic duration. As a reviewer of a previous version of the manuscript pointed out, any spectral reduction we might see as a function of position in the word might be mediated by a tendency of the speaker to increase speaking rate early to late in an utterance. This in turn can lead to reduction of articulatory excursions involved in vowel production and hence to spectral reduction [Lindblom, 1963]. Results of the duration measurements do not bear out this idea. For both talkers, durations of stressed syllables were an inverted V shape, with medial syllables longest. Unstressed syllables showed final lengthening. In the analysis of variance on both talkers with factors stress and position, both main effects and the interaction were highly significant [with the smallest F: $F(2, 174) = 32.01$, $p < 0.0001$].

/a/: Carrier

We expected less evidence of weakening when the trisyllables were produced in a carrier sentence to the extent that weakening, like declination of f_0, is a phrase level, rather than a word level phenomenon. Table 4 provides the descriptive statistics. On stressed syllables of trisyllables, CA showed a progressive decrease only in amplitude [$F(2, 57) = 8.53$]. On unstressed syllables, she showed no declination at all. All acoustic measures on unstressed syllables showed a V pattern like the patterning of F_1 on unstressed /a/ vowels produced in isolation. The pattern was significant for all measures [f_0: $F(2, 117) = 157.6$, amplitude: $F(2, 117) = 5.13$, F_1: $F(2, 117) = 150.7$]. CA's jaw pattern showed a significant *inclination* on unstressed syllables [$F(2, 116) = 107.1$]. On stressed syllables, MV showed monotonic decreases in amplitude [$F(2, 57) = 16.3$] and jaw opening [$F(2, 57) = 72.22$]. F_0 was highest in initial position and equal in medial and final position [$F(2,57) = 6.47$, $p = 0.003$]. MV also showed a monotonic decrease in jaw opening on unstressed syllables [$F(2, 117) = 90.8$]. Except for f_0, other measures on unstressed syllables had minima on the medial syllable [amplitude: $F(2, 117) = 5.16$, $p = 0.007$, F_1: $F(2, 117) = 4.70$, $p = 0.01$]. F_0 increased early to late on unstressed syllables [$F(2, 117) = 47.9$].

Turning to duration measures, as in the no-carrier condition they do not show a pattern of decreased duration early to late. Talker CA shows the same durational pattern as in the no-carrier condition – that is, an inverted V for stressed syllables and final lengthening of unstressed syllables. In an analysis of variance both main effects (position, stress) and the interaction were significant (all $Fs > 11.49$, $p < 0.0001$). Talker MV shows final lengthening of both stressed and unstressed syllables (all $Fs > 26.85$, $p < 0.0001$).

/i/: No Carrier

On trisyllables involving /i/, we can compare only the initial and final syllables. Table 5 provides the means for both talkers in the no-carrier condition. In those utterances, both talkers showed declination of f_0 on stressed syllables [CA: $F(2, 37) = 893.2$, MV: $F(1, 37) = 101.3$]. Only CA showed declination on unstressed syllables [$F(1, 77) = 270.8$]. Both talkers show declination of amplitude on stressed and unstressed syllables [CA: stressed, $F(2, 37) = 195.64$, unstressed, $F(1, 77) = 191.8$; MV: stressed, $F(1, 37) = 9.38$, unstressed, $F(1, 77) = 7.70$].

On F_1, CA, like the speakers in experiment 1, shows an increase in opening for /i/ on the final as compared to the initial syllable [stressed: $F(2, 37) = 118.8$, unstressed $F(1, 77) = 94.7$]. MV shows no significant change in F_1 as a function of position in the word.

In contrast to the talkers in experiment 1, both talkers show increases in F_2 on final as compared to initial syllables. This difference

Table 5. Means and standard deviations (SD) of the various dependent measures of experiment 2, vowel /i/

		Stressed		Unstressed	
		I	F	I	F
No carrier					
CA					
f_0, Hz	mean	242	175	234	179
	SD	6.3	7.7	10.4	18.3
RMS, dB	mean	49.8	42.3	47.7	42.7
	SD	1.6	1.7	1.5	1.7
F_1, Hz	mean	256	316	264	334
	SD	16.7	17.7	30.4	33.8
F_2, Hz	mean	2,626	2,710	2,536	2,705
	SD	65.3	103.3	141.1	236.9
Duration, ms	mean	138	148	87	109
	SD	14.2	24.6	10.0	16.1
MV					
f_0, Hz	mean	132	103	126	126
	SD	12.5	1.2	11.5	15.6
RMS, dB	mean	45.7	42.0	43.4	39.8
	SD	3.8	3.7	4.1	7.1
F_1, Hz	mean	283	281	303	310
	SD	26.7	28.2	11.6	23.3
F_2, Hz	mean	2,129	2,131	1,946	2,155
	SD	47.9	43.9	120.4	227.1
Duration, ms	mean	95	87	62	92
	SD	11.8	19.8	11.9	92
Carrier					
CA					
f_0, Hz	mean	190	208	198	242
	SD	7.9	8.1	7.6	13.2
RMS, dB	mean	44.6	44.1	44.7	44.9
	SD	1.5	1.3	1.8	1.5
F_1, Hz	mean	285	296	316	300
	SD	40.6	41.2	43.7	30.9
F_2, Hz	mean	2,701	2,721	2,599	2,667
	SD	44.5	43.6	92.3	124.3
Duration, ms	mean	125	153	75	126
	SD	8.4	16.0	7.3	18.3
MV					
f_0, Hz	mean	142	142	113	131
	SD	6.7	5.5	4.4	9.1
RMS, dB	mean	48.5	44.4	41.7	43.4
	SD	2.1	2.3	2.4	2.4
F_1, Hz	mean	295	284	305	307
	SD	21.9	30.7	13.3	22.1
F_2, Hz	mean	2,173	2,235	2,000	2,225
	SD	47.9	46.5	118.2	117.4
Duration, ms	mean	99	107	61	68
	SD	9.1	17.1	8.9	6.1

I = Initial; F = final.

is significant for CA on both stressed and unstressed syllables [stressed, $F(1, 37) = 9.24$, unstressed $F(1, 77) = 15.15$]; it is significant for MV only on unstressed syllables [$F(1, 77) = 25.93$].

On the duration measures, CA's final syllables were longer than her initial syllables whether the syllables were stressed or unstressed. The durational difference was greater for the unstressed syllables, however, giving rise to a significant interaction between stress and position [$F(1, 333) = 12.45$, $p = 0.005$] as well as highly significant main effects (both $p < 0.0001$). MV showed final lengthening of unstressed syllables, but his final stressed syllables were a nonsignificant 8 ms shorter than his initial stressed syllables. In the analysis of variance on his durations, both main effects and the interaction were highly significant (all $Fs > 7.94$, $p = 0.006$).

/i/: Carrier

For /i/s produced in a carrier phrase, there is essentially no evidence of declination [excepting a significant decrease in amplitude for MV on stressed syllables, $F(1, 38) = 34.9$]. Here, as in utterances produced in isolation, F_2 of final /i/ was higher than for initial /i/ [CA: stressed $F(1, 38) = 1.02$, NS; unstressed, $F(1, 78) = 7.87$; MV: stressed, $F(1, 38) = 17.56$; unstressed, $F(1, 78) = 73.27$]. Turning to the durational analyses, both talkers produced longer final than initial syllables. For talker CA, the difference was larger for unstressed syllables (41 versus 24 ms); accordingly the interaction of stress and position was significant as well as both main effects (all $Fs > 23.3$, $p < 0.0001$). For talker MV, the position effect was the same in magnitude for stressed and unstressed syllables (and small, 7 ms, in both cases). In the analysis of this talker's data, the main effects of stress and position were highly significant ($Fs > 11.79$, $p = 0.0008$), but the interaction did not approach significance.

Discussion

Overall, the findings on the vowel /a/ of the present experiment replicated those reported in Vayra and Fowler [1987]. For isolated trisyllables, both F_1 and measures of jaw opening decreased monotonically throughout a trisyllable. As in the earlier study, unstressed syllables had their minima medially.

A question the experiment was designed to address was whether this weakening of /a/ in height was a phrase or a word level phenomenon. Results were not wholly consistent; however, overall, they are more consistent with a phrase level interpretation. For CA, declination of F_1 and the jaw is confined to stressed syllables of trisyllables produced in isolation. This is true for F_1 of stressed /a/ of talker MV's utterances; however, measures of jaw opening show a monotonic decline also for trisyllables in a carrier sentence.

A second question the experiment was designed to address was whether declination of supralaryngeal vocalic gestures is coupled to declination of f_0 and amplitude. The answer appears to be 'no'. Whereas declination of f_0 and amplitude occurred over both stressed and unstressed syllables of isolated trisyllables, declination of opening for /a/ occurred only on stressed syllables. Furthermore, in productions involving the vowel /i/, declination of f_0 and amplitude was seen where centralization of the vowel was absent. In general, results on the vowel /i/ did not conform to those seen in experiment 1. Only 1 talker showed decreased closing for stressed /i/ and neither talker showed decreased fronting.

The alternative V pattern for unstressed /a/ seems to have some generality. Both speakers did show the V pattern in jaw and F_1 measures of unstressed /a/ in the no-carrier condition; the same pattern showed up also in F_1 of stressed syllables for CA and in F_1 and amplitude of unstressed syllables for both talkers of trisyllables produced in the carrier sentence, although, in this condition, the jaw measures did not show the same pattern.

Returning to declination itself, we have found some evidence for supralaryngeal weakening over the course of an utterance, but having done so, we acknowledge that we do not know why it occurs when it does. However, if the finding has some generality, then it may be quite important because it reveals that declination is not restricted to the respiratory and laryngeal systems. Possibly it is very general to the articulatory system. Evidence from these talkers suggests, however, that it is largely confined to stressed syllables; perhaps, therefore, the findings index a weakening of the entire mechanism, respiratory, laryngeal and supralaryngeal for stressing a vowel.

Acknowledgements

Research was supported by NINCDS grant 13617 to Haskins Laboratories and a CNR-NATO fellowship to Mario Vayra. We are especially grateful to Cinzia Avesani for her generous contribution of ideas and practical assistance in the conduct of this research. This research was designed jointly by both authors as part of a larger project of comparative studies of Italian and English prosodic structures. In the present work Carol Fowler mainly contributed with measurements and discussion of physiological data and with statistical analysis; Mario Vayra contributed with measurements and discussion of acoustical data and with general discussion.

References

Basbøll, H.: Phonological weight in Italian raddoppiamento fonosintattico. Rivista di Linguistica 1: 5–31 (1989).

Bell-Berti, F.; Krakov, R.: Velar height and sentence length: declination? Spring 1991 Meet. Acoust. Soc. Am., Baltimore, 1991.

Bertinetto, P. M.: A contrastive study on the production and perception of stress by English and Italian speakers; in Conte, Giacalone Ramat, Ramat, Sprache in Kontext, vol. 2, pp. 79–92 (Tübingen 1978).

Bertinetto, P. M.: The perception of stress by Italian speakers. J. Phonet, *8:* 385–395 (1980).

Bertinetto, P. M.: Strutture prosodiche dell'italiano (Accademia della Crusca, Firenze 1981).

Breckenridge, J.: Declination as a phonological process (Bell System, Murray-Hill 1977).

Cohen, A.; Collier, R.; 't Hart, J.: Declination: construct or intrinsic feature of speech pitch? Phonetica *39:* 254–273 (1982).

Cooper, W.; Sorensen, J.: Fundamental frequency in sentence production (Springer, New York 1981).

Farnetani, E.; Kori, S.: Lexical stress in spoken sentences: a study on duration and vowel formant pattern. Quaderni del Centro di studio per le ricerche di fonetica del C.N.R. *1:* 106–133 (1982).

Fava, E.; Magno Caldognetto, E.: Studio sperimentale delle caratteristiche elettroacustiche delle vocali toniche e atone in bisillabi italiani; in Simone, Vignuzzi, Ruggiero, Studi di fonetica e fonologia, pp. 35–80 (Bulzoni, Roma 1976).

Ferrero, F.: Caratteristiche acustiche dei fonemi vocalici italiani. Parole e Metodi *3:* 9–32 (1972).

Fowler, C. A.: An event approach to the study of speech perception from a direct-realist perspective. J. Phonet. *14:* 3–28 (1986).

Fowler, C. A.; Smith, M.: Speech perception as 'vector analysis': an approach to the problems of segmentation and invariance; in Perkell, Klatt, Invariance and variability of speech processes, pp. 123–136 (Lawrence Erlbaum, Hillsdale 1986).

Gelfer, C.: A simultaneous physiological and acoustic study of fundamental frequency declination; thesis City University of New York (1987).

Gelfer, C.; Harris, K.. Baer, T.: Controlled variables in sentence intonation; in Baer, Sasaki, Harris, Laryngeal function in phonation and respiration, pp. 422–435 (College-Hill, Boston 1987).

Gulisano, V.: Physiological signal processing system: user's manual (Haskins Laboratories, New Haven 1982).

't Hart, J.; Collier, R.; Cohen, A.: A perceptual study of intonation (Cambridge University Press, Cambridge 1990).

ILS Command Reference Guide V 6.0 (Signal Technology, Goleta 1986).

Kay, B. A.; Munhall, K. G.; Vatikiotis-Bateson, E. V.; Kelso, J. A. S.: A note on processing kinematic data: sampling, filtering and differentiation. Haskins Lab. Status Rep. Speech Res., SR *81,* pp. 291–303 (Haskins Laboratories, New Haven 1985).

Koopmans-van Beinum, F.: Vowel contrast reduction: an acoustical and perceptual study of Dutch vowels in various speech conditions; thesis University of Amsterdam (1980).

Koopmans-van Beinum, F.: Systematics in vowel systems; in van den Broecke, van Heuven, Zonneveld, Sound structures. Studies for Antonie Cohen, pp. 159–171 (Foris, Dordrecht 1983).

Lieberman, P.; Katz, W.; Jongman, A.; Zimmerman, R.; Miller, M.: Measures of the sentence intonation of read and spontaneous speech in American English. J. acoust. Soc. Am. *77*: 649–657 (1985).

Lindblom, B.: A spectrographic study of vowel reduction. J. acoust. Soc. Am. *35*: 1773–1781 (1963).

Maeda, S.: A characterization of American English intonation; thesis Massachusetts Institute of Technology (1976).

Marotta, G.: Modelli e misure ritmiche (Zanichelli, Bologna 1985).

Os, E. den: Rhythm and tempo of Dutch and Italian: a contrastive study; thesis Utrecht University (1988).

Os, E. den: Vowel reduction in Italian and Dutch. P. R. I. P. V. *10:* 3–12 (Institute of Phonetics Univ. Utrecht 1985).

Pierrehumbert, J.: The perception of fundamental frequency declination. J. acoust. Soc. Am. *66*: 363–69 (1979).

Silverman, K.: The structure and processing of fundamental frequency contours; thesis University of Cambridge (1987).

Terken, J.: Fundamental frequency and perceived prominence of accented syllables. J. acoust. Soc. Am. *89*: 1768–1776 (1991).

Umeda, N.: 'F_0 declination' is situation dependent. J. Phonet. *10*: 279–290 (1982).

Vayra, M.; Fowler, C. A.: The word level interplay of stress, coarticulation, vowel height and vowel position in Italian. Proc. 11th Int. Congr. Phonet. Sci., vol. 4, pp. 24–27 (Academy of Sciences of the Estonian SSR, Tallin 1987).

Vayra, M.; Fowler, C. A.; Avesani, C.: Word-level coarticulation and shortening in Italian and English speech. Studi di Grammatica Italiana *13*: 249–269 (1987); also in Haskins Lab. Status Rep. Speech Res., SR., *91*, pp. 75–89 (Haskins Laboratories, New Haven 1987).

Vincent, N.: Non-linear phonology in diachronic perspective: stress and word-structure in Latin and Italian; in Bertinetto, Loporcaro: Certamen phonologicum. Papers from the 1987 Cortona Phonology Meeting, pp. 421–432 (Rosenberg & Sellier, Torino 1988).