Journal of Phonetics (1998) **26**, 45–69

Rate effects on French intonation: prosodic organization and phonetic realization

Cécile Fougeron*[†] and Sun-Ah Jun*

**Department of Linguistics, UCLA, 405 Hilgard, Ave., Los Angeles, CA 90095, U.S.A.*
[†]*Paris III-Sorbonne Nouvelle, CNRS URA-1027, France*

Received 23rd October 1996, and in revised form 3rd October 1997

This study shows that an increase in speech rate affects French intonation in both the phonetic realization of the f_0 contour and the prosodic organization of a text. The effect of rate was found to vary according to speaker and position of the speech material within the text. For two of the three speakers, an increase in rate induced a reduction in pitch range and in pitch displacements, with greater lowering of f_0 maxima than of f_0 minima. For one of these two speakers, modifications in the shape of the f_0 contour were found only in the first part of the text. For the third speaker, an increase in rate induced a reduction of pitch range only in the second part of the text, and a reduction of pitch displacements was obtained by raising the f_0 minima while keeping the maxima constant. The prosodic organization was also modified at fast rate. Speakers reduced the number of phrases by deleting or reducing the strength of prosodic boundaries. Simplification of the tonal contour was also found at fast rate with the failure to realize certain underlying tones. These results are compared to rate-based variation in the kinematics of other articulators. Evidence for a saturation effect is discussed with regard to the variation observed in the compressibility of f_0 movements at fast rate. A model of the articulation of intonation is proposed which assumes that speakers use similar strategies in controlling laryngeal/subglottal articulators and other articulators.

1. Introduction

The effect of speaking rate has mostly been observed at the segmental level, with a focus on modifications to the temporal and spatial characteristics of speech. In general, three kinematic variables have been observed to vary when speaking rate is increased. Speakers may reduce the spatial magnitude of articulatory movements, described as "target undershoot" (e.g., Lindblom, 1963, 1964; Kent & Moll, 1972; Gay, 1981). They may adjust the speed of transitions between successive targets by increasing the velocity of their movements (e.g., Abbs, 1973; Kuehn & Moll, 1976). Finally, they may increase the overlap between successive articulatory gestures by modifying their phasing (e.g., Engstrand, 1988; Munhall & Löfqvist, 1992; Krakow, 1993). These three variables are

Address correspondence to S.-A. Jun, UCLA, Department of Linguistics, 3125 Campbell Hall, Box 951543, Los Angeles, CA 90095-1543, U.S.A.

0095–4470/98/010045 + 25 $25.00/0/jp970062

not mutually exclusive and can interact with each other in order to shorten articulation time. In fact, several studies have shown that speakers vary in how these three parameters are employed as speaking rate increases (e.g., Kuehn, 1976; Kuehn & Moll, 1976; Ostry & Munhall, 1985).

The effect of speaking rate in the suprasegmental domain has also been studied, but most work focuses on variation in duration (e.g., for French, Bartkova, 1991; Keller & Zellner, 1995). The effect of rate on intonation has rarely been studied. Recently, Caspers and van Heuven examined in detail the effects of rate on pitch movements in Dutch (Caspers & van Heuven, 1991, 1993, 1995; Caspers, 1994). They looked at variation in the number and type of pitch events induced by an increase in speech rate. They found that the tonal configurations used were similar at normal and fast rates. They also found that the number of melodic boundaries was reduced at fast speech rate while the number of pitch accents remained constant. Reducing the number of boundaries means reducing the number of phrases. This reduction in the number of phrases at fast rate has also been discussed for French by Vaissière (1983) and for Korean by Jun (1993). In Jun's study of Korean, as the rate increased, the number of accentual phrases (a phrase with no lexical pitch accent and smaller than an intonational phrase) decreased by approximately 24%, averaged across five speakers. Interestingly, it seems that not all phrase boundaries are equally affected at fast rate: Caspers and van Heuven (1991, 1995) found that in Dutch, boundary marking pitch movements were more likely to be deleted in the case of optional intonational phrase boundaries (i.e., end of NP, before an embedded sentence, before PP).

Variation in the phonetic realization of the fundamental frequency (f_0) contour has also been observed at fast rate. Kohler (1983) found that in German there was a general raising of the phonatory level, both in frequency and intensity, at fast rate. Also, the level of the f_0 peak (the highest f_0 value in a sentence) was maintained, whereas the level of the f_0 valleys was raised at fast rate. In consequence, the pitch displacements between peaks and valleys were reduced. A different pattern was found in Dutch by Caspers and van Heuven: levels of both peaks and valleys were raised at fast rate. As for pitch displacements, they observed variation between speakers as they raised their peaks and valleys in different proportions. One speaker raised his peaks more than his valleys, resulting in an increase in pitch displacements, while the other speaker raised his valleys more than his peaks, resulting in a decrease in pitch displacements.

In sum, it seems consistent across studies that an increase in rate leads to a reduction in the number of phrases. However, the effect of rate on pitch displacements and on the level of peaks and valleys seems to vary.

In this study, we investigated the effect of an increase in speech rate on the phonetic realization as well as the phonological structure of French intonation. *By phonetic realization*, we mean the shape of the pitch contour: the level of peaks and valleys, pitch displacements, velocities of pitch movements, and pitch range (highest minus lowest f_0 value). By *phonological structure*, we mean the prosodic organization of an utterance, which is cued by its prosodic phrasing (i.e., the number and types of boundaries), the type of tones used, and its tonal pattern. Both phonetic and phonological representations were examined because the intonation of an utterance is the variation in fundamental frequency and, at the same time, the reflection of an abstract prosodic organization.

We examined the effect of rate on the intonation pattern of a text instead of on isolated sentences. We expect that, because the discourse structure is richer in a text than in isolated sentences, the intonation pattern of a text will provide phrases of various sizes with a larger inventory of boundary tone types, as well as more variation in pitch range than that of an isolated sentence (Hirschberg & Pierrehumbert, 1986). Furthermore, we divided the text into two parts at a narrative break, and compared the modifications made in these two parts. Since local articulation rate has been found to vary within long stretches of speech (Brubaker, 1972; Miller, Grosjean & Lomanto, 1984) and depending on prosody and discourse structure (Crystal & House, 1990; Koopmans-van Beinum & van Donzel, 1996), we expect that the effect of speech rate on f_0 may also vary throughout the text, depending on discourse structure and location within a text. Speakers may have to use a different strategy towards the end of the text where f_0 declination induces a lowering of the f_0 top-line ("plateau") and a reduction of pitch range (Vaissière, 1983). We also hypothesize that important boundaries in the discourse structure will vary less than other boundaries, as found by Caspers & van Heuven (1991, 1995).

Since speakers have been found to vary in their strategies for increasing speech rate, we tested for speaker variation both in terms of phonetic realization and phonological organization of intonation by comparing the speech of three French speakers.

Finally, to understand the global mechanism involved in speech production at fast rate, we compared the adjustments made in the intonation domain with those described in the literature for the segmental domain. Therefore, we tried to use terminology similar to that used in articulatory studies. We interpret the f_0 contour in terms of *pitch targets* (level of f_0 peak and valley), *pitch displacements* (pitch rise and fall), and *average pitch movement velocity* (slope of pitch rise and fall). In this approach, we assume that a pitch contour consists of a sequence of pitch targets and each target has an articulatory goal, laryngeal and/or subglottal. That is, we hypothesize that there is an active articulation of intonation, and that its adjustments to increase rate are governed by the same kind of mechanism that controls other articulators.

2. Method

2.1. *Subjects and speech material*

Three Parisian French speakers (two female and one male) in their twenties participated in the experiment. Subjects were asked to read the text "La bise et le soleil" ("The North Wind and the Sun"), given in Table I, at self-selected normal and fast rates. The text was presented to the speakers in normal orthography including capitals and punctuation marks. Speakers were told to read the story in a lively manner, but no special instruction was given regarding the phrasing of the text. The recordings were made in a sound booth, and the speakers were asked to read the text three times at normal rate followed by three times at fast rate. Each time a speaker misread some part of the text, she or he had to start again from the beginning. No speech "error" or hesitation breaks were therefore included in the recordings.

The text was chosen for its narrative aspects (a tale) as well as for its syntactically complex construction, which yielded a large variety of phrases. To examine whether speakers use different strategies to increase rate depending on position in the text, the text was divided into two parts at a semantic and narrative juncture. The first part consists of two long sentences and is the introduction to the tale, and the second part consists of four sentences and is the development and conclusion of the tale.

TABLE I. Text "La bise et le soleil". The numbers in parentheses correspond to the boundary codes used in Figs. 5 and 6. The break between the two parts and the codes in parentheses were not written in the version read by the subjects, but punctuation and capitals were. Word-for-word translation into English is given below the French text

La Bise et le Soleil

First part: La bise (1) et le soleil (2) se disputaient (3), chacun (4) assurant (5) qu'il était (6) le plus fort (7). Quand ils ont vu (8) un voyageur (9) qui s'avançait (10), enveloppé (11) dans son manteau (12), ils sont tombés (13) d'accord (14) que celui (15) qui arriverait (16) le premier (17) à le lui faire (18) ôter (19) serait (20) regardé (21) comme le plus fort (22).

Second part: Alors (23), la bise (24) s'est mise à souffler (25) de toutes ses forces (26), mais plus elle soufflait (27), plus le voyageur (28) serrait son manteau (29) autour de lui (30). Finalement (31), elle renonça (32) à le lui faire ôter (33). Alors (34), le soleil (35) commença à briller (36) et au bout d'un moment (37) le voyageur, réchauffé (38), ôta son manteau (39). Ainsi (40), la bise (41) dut reconnaître (42) que le soleil (43) était le plus fort (44).

The North Wind and the Sun

First part: The North Wind (1) and the Sun (2) were arguing (3), each (4) claiming (5) that he was (6) the stronger (7). When they saw (8) a traveler (9) who appeared (10), wrapped (11) in his coat (12), they made (13) an agreement (14) that whoever (15) should succeed (16) first (17) in making him (18) take it off (19) would be (20) considered (21) the stronger (22).

Second part: Then (23), the North Wind (24) began to blow (25) with all its might (26), but the more it blew (27), the more the traveler (28) wrapped his coat (29) around himself (30). Finally (31), it gave up (32) in making him take it off (33). Then (34), the Sun (35) began to shine (36) and after a moment (37) the traveler, all warmed up (38), took off his coat (39). So (40), the North Wind (41) had to acknowledge (42) that the Sun (43) was the stronger (44).

2.2. *Measurements*

In this study, the term "f_0" and "pitch" are used interchangeably. Pitch tracks of the readings were extracted and analyzed using Entropic Research Laboratory's XWAVES + speech analysis software. Several acoustic measurements were taken for each tonal pattern. These are schematically illustrated in Fig. 1. For each f_0 peak and valley, the pitch level at the peak (hereafter called f_0 *maximum*) and at f_0 rising onset and falling offset (both points hereafter called f_0 *minima*) were collected. When the valley or peak occurred during an obstruent, f_0 values were taken in the nearest vowel. The difference in Hz between the minima and the maxima gives the magnitude of the pitch movements, which we will call pitch *displacements* in frequency (Hz) between these pitch targets [equivalent to the "excursion size" of the rise and fall reported by Caspers (1994)]. The *average velocity* (rate of change) of the pitch movement was calculated by dividing the displacement (in Hz) by the *transition* time between the minimum and maximum (in ms). Rising or falling pitch movements were measured only between adjacent minimum and maximum values. Thus, when f_0 stays low or plateaus, no pitch movements were taken. For each speaker, *pitch range* was also calculated by taking the difference between the highest f_0 value and an "estimated" lowest f_0 value. Since the lowest f_0 values are likely to fall on the word "fort" (major L boundaries) and since f_0 perturbations might affect the reliability of the measures on this word, we estimated the lowest f_0 value as the lowest value reached by more than 6 data points (2 "fort" in the text by 3 repetitions). Attention must be directed to the fact that in our terms *pitch range* is not the *averaged*

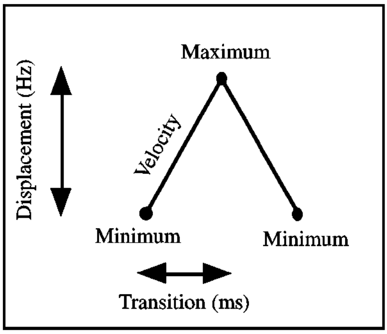

Figure 1. Schematic drawing of measurements taken from an f_0 contour.

range of f_0 displacement (as in Caspers, 1994), but the *maximal* range of f_0 variations in the text for a particular speaker. All measurements were compared at fast and normal rate (averaged across three repetitions with the same prosodic phrasing) in order to evaluate modifications in the shape of the f_0 contour at fast rate. Furthermore, to assess the relative rate within and across speakers, the number of pauses and pause duration within the whole text, as well as duration of each syllable, were measured.

Next, to examine modifications in the prosodic organization, the prosodic structure of the text was qualitatively compared at fast and normal rate. The model of French intonation developed by Jun & Fougeron (1995) was used for the labeling of each pitch event. Following Pierrehumbert and others (Liberman, 1975; Pierrehumbert, 1980; Beckman & Pierrehumbert, 1986), we assumed in this model that a tune is composed of a sequence of underlying H and L tones, and that the intonational structure of a sentence is hierarchically organized. Two prosodic levels, and therefore two types of boundaries, are defined for French. The lowest tonally defined prosodic level is the Accentual Phrase (AP). This level is marked by a minor continuation rise (Delattre, 1966) and a small final lengthening. This level is higher than the Tonal Unit of Hirst & Di Cristo (1984, in press), and roughly corresponds to the "prosodic word" (Vaissière, 1992) or Intonation Group (Mertens, 1993), or to the Intonème Mineur (Rossi, 1985) in French. The Accentual Phrase has the underlying tonal representation /L Hi L H*/. The final high tone (H*)[1] is realized on the phrase-final full syllable ("accent primaire") and the initial high tone (Hi) is realized on one of the first three syllables of the phrase (related to the notion of "accent initial, secondaire", "ictus mélodique", "initial rise"). Each L tone is realized on the syllable preceding the H-toned syllable. Values of f_0 on other syllables are considered to be interpolated between adjacent pitch targets. An example of the [L Hi L H*] tonal realization of an AP is illustrated on "serait regardé" in Fig. 2 (top panel).

However, the four underlying tones of an AP are not always realized on the surface. For example, the Hi is often not realized, depending on rhythmic or pragmatic factors (cf. Rossi, 1985; Pasdeloup, 1990), and the AP surfaces with a tonal pattern [LLH*] as in

[1] The H* here differs from that of English since French H* is a phrasal pitch accent linked to a phrasal accented syllable, not to a lexically stressed syllable. The diacritic "*" is used to show that the tone is associated to one particular syllable with long duration. We assume that French has accentable syllables (like the full final syllable of lexical words) and that this pitch accent is realized on one of these syllables.

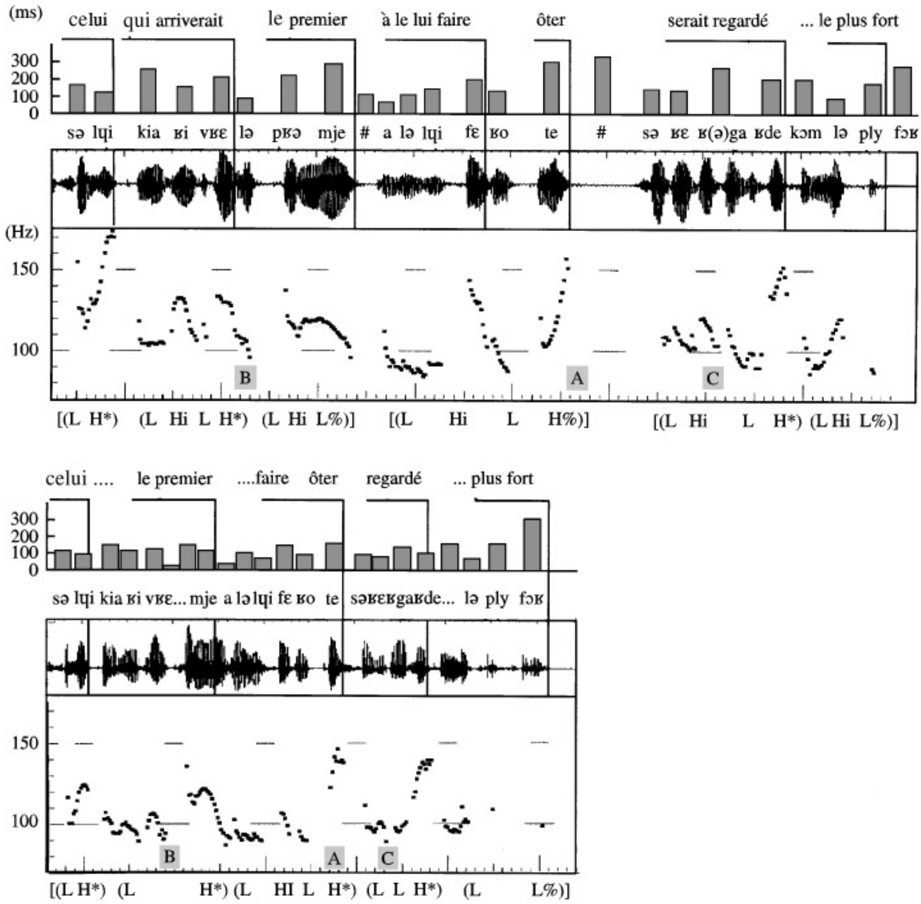

Figure 2. Example of "qualitative" reduction found in the prosodic organization. The upper graph shows the waveform, f_0-track, and prosodic labeling of the sequence "celui qui arriverait le premier à le lui faire ôter serait regardé comme le plus fort" produced by speaker 3M at normal rate. The lower graph shows the same sequence produced by the same speaker at fast rate. Duration of each syllable (or a sequence of segments that can easily segmented on the waveform) is indicated in the bar graph above the waveform and the IPA transcription. The letters A, B, C refer to specific examples of reduction discussed in the text. For the prosodic labeling: [ ] = IP boundaries, ( ) = AP boundaries, L = low tone, Hi = initial High tone, H* = phrasal pitch accent for the primary stressed syllable in the AP, % = IP boundary tone.

"serait regardé" in Fig. 2 (bottom panel). Other surface forms have been observed, depending on the adjacent tonal sequence: [HiLH*], [LHiH*], [LHiL*] (for more detail on the phonetic realization rules involved, see Jun & Fougeron, 1995).

A prosodic level higher than an AP, and the highest level in this model, is the Intonational Phrase (IP). The IP is marked by a major continuation rise or a major final fall. This level is also marked by a large final lengthening (e.g., Hirst & Di Cristo, 1984; Pasdeloup, 1990) and is optionally followed by a pause. This level in French roughly corresponds to the Intonème Majeur of Rossi (1985) and the Unité Intonative of Hirst

& Di Cristo (1984, in press). Following the notation developed for English by Pierrehumbert, we will note the right tonal boundary of the IP by a H% or an L% tone. In Fig. 2, the three IPs are delimited by either a low boundary tone (L%), at the end of "premier" and "fort", or a high boundary tone (H%), at the end of "ôter". All of the three IPs are delimited by a pause.

Following this model, f_0 points relevant to the prosodic organization (L, Hi, H*, H%, L% tones) were labeled by hand for each repetition of the text by examining the f_0 curves (slope of falling and rising, tone-syllable association, etc.), noting duration and loudness of the syllables, and listening to the speech for the auditory impression of juncture. The transcription was done first by the first author, a native speaker of French, and then by the second author at a separate time. In cases where the second author did not agree with the first author, the first author was asked to reexamine the case. In the few cases of further disagreement (mostly about the presence or absence of an AP boundary) we inclined to the native speaker's judgement of juncture.

Both the phrasing of the text and the tonal realization of the APs were compared between normal and fast rates. Note that the difference in the number of accents/boundaries as a function of position in the text is not an issue in this study; only the differential effects of rate are. Fig. 2 gives an example of the types of qualitative reduction found in the prosodic organization. The upper graph shows the waveform and f_0 track of the sequence "celui qui arriverait le premier à le lui faire ôter serait regardé comme le plus fort" produced by speaker 3M at normal rate. The lower graph shows the same sequence produced by the same speaker at fast rate. Duration of each syllable (or a sequence of segments that can be easily segmented on the waveform) is given in the bar graphs above the waveforms. Prosodic labeling is given below the f_0 track with each label aligned with the tonal event. The letters A, B, and C refer to specific examples of reduction. In (A), the IP-boundary at the end of "… le lui faire ôter" is reduced to an AP boundary at fast rate. This kind of reduction will be denoted as [IP ⇒ AP] (in Table IV below). In this case, the sentence that consists of three IPs at normal rate is produced as one IP at fast rate. Similarly, in (B) the two APs at normal rate, "qui arriverait" and "le premier", are grouped in one AP at fast rate. Here, the reduction is accomplished by deletion of the AP-final boundary (denoted [AP ⇒ Ø] in Table IV). (C) shows an example of a change in the realization of the underlying tonal pattern of an AP. The initial Hi tone in the AP "serait regardé" is not realized at fast rate (denoted [Hi ⇒ (Hi)] in Table IV).

3. Results

3.1. *Rate characteristics*

In order to assess relative speaking rate within and across speakers, durational characteristics of the text were compared. Table II lists the differences found between the two rates for the three speakers in terms of total duration of the text, number of pauses, average duration of pauses, articulation rate (excluding pause), speaking rate (including pause), and average syllable durations and their coefficient of variation. Results are given for the two parts of the text, allowing comparison within and across speakers. The normal articulation rate ranges from 5.2 to 6 syll/s across speakers and is comparable to that found in previous studies of conversational speech in French: 5.73 syll/s by Malécot, Johnson & Kizziar (1972) and 5.29 syll/s by Grosjean & Deschamps (1975).

TABLE II. Rate characteristics of the speakers averaged over three repetitions at normal (N) and fast (F) rates. Total duration of the part (in s), number of pauses, mean pause duration (in ms), articulation rate (in syll/s, pause excluded), increase in articulation rate from normal to fast rate (in syll/s), speaking rate (in syll/s, pause included), and mean syllable duration (in ms) with coefficient of variation (in %)

	Speaker 1F				Speaker 2F				Speaker 3M			
	1st part		2nd part		1st part		2nd part		1st part		2nd part	
	N	F	N	F	N	F	N	F	N	F	N	F
Total duration	16.0	10.7	21.3	14.7	15.4	11.5	20.3	14.8	16.5	9.1	21.9	12.6
Mean # pauses	6.0	3.3	7.3	4.7	7	5.7	14	11	8.3	2.7	13	5.7
Pause duration	639	334	579	400	495	224	385	199	478	277	444	239
Articulation rate	5.7	7.2	5.2	6.9	6	7.2	5.9	7.2	5.5	8.2	5.6	8
Increase	1.5		1.7		1.2		1.3		2.7		2.4	
Speaking rate	4.3	6.4	4.2	6.1	4.6	6.4	4.5	6.1	4.2	7.6	4.1	7.2
Syll. duration	177	140	192	144	168	139	169	137	182	122	181	126
Variation coeff.	0.4	0.4	0.4	0.3	0.4	0.3	0.4	0.3	0.4	0.4	0.4	0.4

The self-selected fast rate varies across speakers. The male speaker (3M) has the largest rate increase (2.4–2.7 syll/s faster) and the second female speaker (2F) has the smallest rate increase (1.2–1.3 syll/s faster); the first female speaker (1F) has a medium increase (1.5–1.7 syll/s faster). For all speakers, the average syllable duration is significantly shorter at fast compared to normal rate. Syllable duration was reduced by 23% overall for speaker 1F ($t = 10.7$, $p < 0.0001$; $\mathrm{df} = 945$), by 17% for speaker 2F($t = 8.4$, $p < 0.0001$; $\mathrm{df} = 970$), and by 32% for speaker 3M ($t = 13.6$, $p < 0.0001$; $\mathrm{df} = 940$).

Since speech rate may vary depending on position in the text and discourse structure, and since speakers may not use a consistent strategy to increase the speaking rate throughout the whole text, we examined articulation rate and syllable duration in two successive parts of the text at fast and normal rates. Table II shows that speaker 2F has a similar articulation rate increase from normal to fast rate in the two parts. For speaker 1F, the rate increase is larger in the second part of the text; while for speaker 3M, the rate increase is smaller in the second part of the text. However, these differences in the increase of articulation rate depending on position in the text are quite small, and when looking at the effect of rate on syllable duration no significant interaction of rate with text part is found ($F(1, 943) = 1.56$, $p = 0.21$ for 1F; $\mathrm{F}(1, 968) = 0.07$, $p = 0.79$ for 2F; $F(1, 938) = 0.32$, $p = 0.57$ for 3M).

Since substantial variation in articulation rate can be averaged out when articulation rate is measured over large stretches of speech (Miller *et al.*, 1984), we calculated the articulation rate for each "chunk" of the continuous speech delimited by a pause (similar to Miller *et al.*'s (1984) "run of pause-free speech"). This allowed us to test whether the increase of rate is constant across smaller units of speech and whether this varies depending on the position of the unit in the text. Since these chunks correspond to IPs and are delimited by a pause at normal rate but not always at fast rate, we called them "C" to avoid confusion. The number and the content of "C" are not always the same across speakers but depended on their phrasing of the text. Fig. 3 shows the articulation rate per "C" at normal and fast rates for speaker 1F. The other two speakers show the same pattern. When the articulation rate between successive C's is compared at both normal and fast rates, no systematic trend is found depending on the position in the text. Some individual C's may seem faster or slower than others, but this variation across C's is not a function of their position. In fact, no correlation is found between the articulation rate of each chunk and the linear position of the chunk in the text (for the three speakers, at both rates, r^2 was never higher than 0.09). These results do not support Brubaker's (1972) observation of a progressive increase of rate within paragraphs, nor Koopmans-van Beinum & van Donzel's (1996) observation that speakers slow down at the start of a new paragraph and speed up at the end of a paragraph. On the contrary, our results corroborate Miller *et al.*'s (1984) observation that during a given speech interval, articulation rate does not increase or decrease gradually, but changes course a number of times. In sum, it seems that no major variation in rate increase in found throughout the text for the three speakers.

3.2. *Phonetic modification in the shape of f_0 contour*

3.2.1. *Pitch range*

The differences in pitch range between fast and normal rates are presented in Fig. 4 and Table III for the first and the second parts of the text. In the first part of the text, pitch

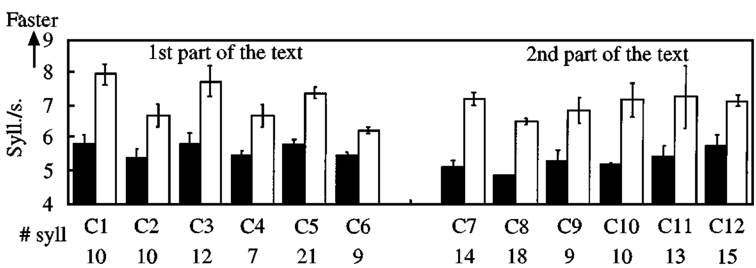

Figure 3. Articulation rate (syllables/s for each "chunk" (C)) in the text at normal (■) and fast rate (□) rates; see text for the definition of C. Below each C, the number of syllables within each C is given. Data for speaker 1F.

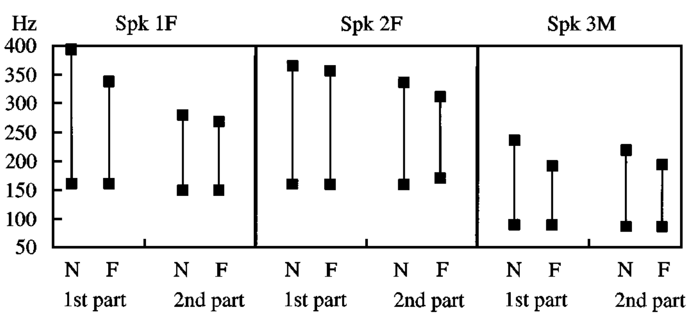

Figure 4. Pitch range (highest–lowest f_0 value) at normal (N) and fast (F) rates, in the first and second parts of the text for the three speakers.

TABLE III. Reduction in the shape of f_0 contour at fast speech rate. Mean values at norm (N) and fast (F) rates and % of reduction at fast rate compared to norm rate (100(N − F/N). Velocity in Hz/ms, other measures in Hz

	Speaker 1F				Speaker 2F				Speaker 3M			
	1st part		2nd part		1st part		2nd part		1st part		2nd part	
	N	F	N	F	N	F	N	F	N	F	N	F
f_0 maximum	262	233	215	208	275	273	254	252	165	143	145	123
	11%		3%		1%		1%		13%		15%	
f_0 minimum	192	178	173	166	205	213	192	197	106	101	100	95
	7%		4%		−4%		−3%		5%		5%	
Rising Displacement	65	55	40	41	66	55	57	49	60	40	46	27
	15%		−3%		17%		14%		33%		41%	
Falling Displacement	65	56	43	41	68	58	60	53	61	45	47	28
	14%		5%		15%		12%		26%		40%	
Rising Velocity	0.35	0.34	0.26	0.27	0.41	0.41	0.37	0.37	0.29	0.24	0.27	0.19
	3%		−4%		0%		0%		17%		30%	
Falling Velocity	0.38	0.36	0.21	0.23	0.34	0.32	0.22	0.25	0.30	0.28	0.18	0.19
	5%		−10%		6%		−14%		7%		−6%	
Pitch range	233	178	128	118	204	196	177	143	147	103	132	108
	24%		8%		4%		19%		30%		18%	

range is considerably reduced at fast rate for speakers 1F and 3M (24% and 30%, respectively) by lowering the highest f_0 value and keeping the lowest f_0 value the same. In contrast, for speaker 2F, the highest and lowest f_0 values are similar at both rates. Therefore, this speaker shows very little reduction of pitch range at fast rate in the first part (4%). In the second part of the text, the pattern is different. Speaker 1F has very little reduction of pitch range at fast rate (8%) compared to that in the first part. Speaker 3M has also less reduction of pitch range than in the first part, although the reduction at fast rate is still large (18%). Speaker 2F, on the other hand, shows much more reduction at fast rate in the second part of the text (19%) than in the first part. This reduction results from a lowering of the highest f_0 value and a raising of the lowest f_0 value at fast rate.

3.2.2. *Pitch targets* (*maxima and minima*) *and pitch displacements*

Values of f_0 at fast and normal rates, and the reduction in the shape of f_0 contour at fast speech rate are shown in Table III for the three speakers. In the first part of the text, the rate increase induces an overall lowering of both types of pitch targets for speakers 1F and 3M. Maxima in f_0 are lower at fast rate than at normal rate, and this lowering is statistically significant ($F(1,107) = 12.7$, $p < 0.001$ for 1F; $F(1, 120) = 17.2$, $p < 0.0001$ for 3M). Fig. 5(a) illustrates, for speaker 1F, the difference in pitch level for the f_0 maxima in the first part of the text at normal (black bar) and fast (white bar) rates. For each peak, the number given on the abscissa corresponds to the boundary code written in parentheses after each prosodic phrase in the text given in Table I. For this speaker, all f_0

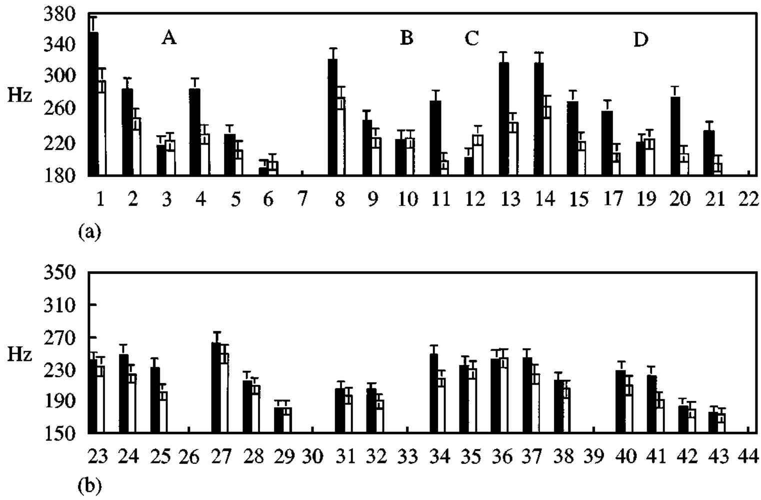

Figure 5. Maxima of f_0 in the first (a) and second (b) parts of the text for speaker 1F. For each maximum, the number on the abscissa corresponds to the boundary code given in Table I. For each number, f_0 maxima are given at normal (■) and fast (□) rates. Letters A, B, C, and D indicate the major prosodic boundaries that do not reduce at fast rate.

maxima except the peaks written in Fig. 5 as A to D are lowered at fast rate. These four f_0 peaks correspond to the major IP boundaries with a H boundary tone (H%) in the text. These are "…se disputaient" (A), "… qui s'avançait" (B), "… son manteau" (C), and "… faire ôter" (D). For speaker 3M (not illustrated here), all f_0 maxima but one are lowered at fast rate. For these two speakers, and in this first part of the text, f_0 minima are also significantly lowered at fast rate ($F(1, 166) = 19.9, p < 0.0001$ for 1F; $F(1, 172) = 5.4$, $p = 0.02$ for 3M). This lowering (7% for 1F and 5% for 3M) is smaller than that of the maxima (11% for 1F and 13% for 3M). As a consequence, there is a noticeable reduction of pitch displacements, for both rising f_0 displacement (15% for 1F and 33% for 3M) and falling f_0 displacement (14% for 1F and 26% for 3M). However, this reduction of pitch displacement is significant only for speaker 3M ($F(1, 112) = 11.7$, $p < 0.0001$ rising; $F(1, 113) = 9.9$, $p < 0.001$ falling). Speaker 2F, on the other hand, shows a different strategy in the reduction of f_0 maxima and minima at fast rate: f_0 maxima are not reduced and f_0 minima are raised. Fig. 6 illustrates this pattern for f_0 maxima (6a) and f_0 minima (6b) at fast rate (white bars) *vs.* normal rate (black bars) in the first part of the text. It can be seen that f_0 maxima often have either the same height at both rates or are inconsistently raised or lowered at fast rate, while almost all f_0 minima are slightly raised at fast rate. This raising of f_0 minima at fast rate is small (4%) but nearly significant ($F(1, 195) = 3.9$, $p = 0.05$). As a consequence, this speaker also reduces her overall pitch displacements at fast rate, but with a different strategy: she raises her f_0 minima whereas the other two speakers lower their f_0 maxima. Reduction of pitch displacement for this speaker is similar for rising (17%) and falling (15%) movements, but is significant only for the former ($F(1, 147) = 4.3$, $p < 0.04$)

In the second part of the text, speakers 3M and 2F keep the same strategy as in the first part of the text (lowering significantly both pitch targets for 3M, rising slightly the

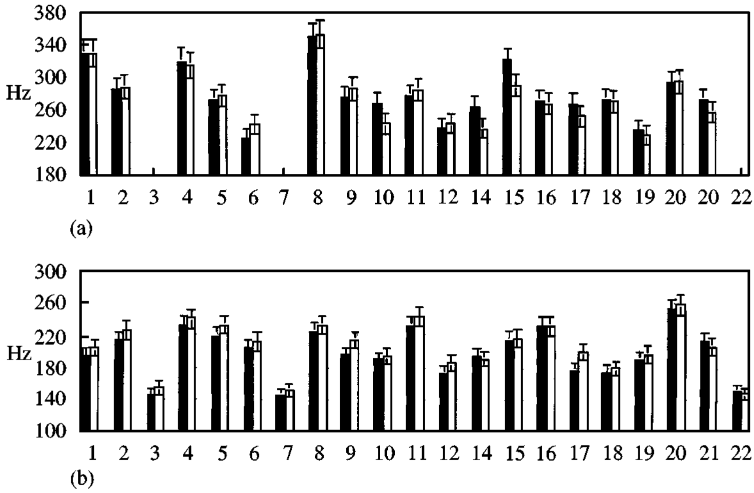

Figure 6. Maxima of f_0 (a) and minima (b) in the first part of the text for speaker 2F. The numbers on the abscissa correspond to the boundary codes given in Table I; f_0 value for normal (■) and fast (□) rates.

minima for 2F). By contrast, for speaker 1F, the effects of rate increase on pitch targets and displacements differed from the first to the second part of the text. Fig. 5(b) illustrates the effect of rate on the f_0 maxima in the second part of the text for this speaker. The reduction of the maxima at fast rate is on average much smaller (3%) than that observed in the first part (11%), and is not significant ($F(1, 163) = 3.2$, $p = 0.07$). The reduction of the minima, not illustrated here but given in Table III, is smaller in the second part (4%) but is still significant ($F(1, 207) = 10.5$, $p = 0.001$). As a consequence, pitch displacement shows little or no reduction in this part.

3.2.3. *Pitch movement average velocity*

An increase of rate is found to have various effects on the average velocity of the pitch movements: it may increase, decrease, or remain the same. In most of the cases the average velocity of pitch movements is slightly reduced at fast rate (indicated in Table III by a positive %). This reduction is significant only in one case (rising displacement in the second part of the text for 3M). In fewer cases, we found an increase of movement velocity at fast rate (indicated by a negative % in Table III), but the increase is generally small and is never significant. Thus, it is never the case that f_0 movements are significantly faster and steeper at fast rate than at normal rate.

3.3. *Phonological modification in prosodic structure of the text*

Regarding the prosodic organization of the text, we observed modification both in the phrasing of the text and in the realization of the underlying tonal pattern at fast rate. The results obtained in the two parts of the text are given in Table IV for each speaker. Descriptive examples of the three types of organizational modification observed were given in Fig. 2 and explained in Section 2.2.

TABLE IV. Modification in prosodic organization at fast speech rate. Percentage of reduction per type of prosodic reduction by speaker. In parentheses, number of boundaries (accent) reduced at fast rate/number of boundaries (accent) at normal rate. IP = intonational phrase, AP = accentual phrase, Hi = initial high tone

	Speaker 1F		Speaker 2F		Speaker 3M	
Phrasing	1st part	2nd part	1st part	2nd part	1st part	2nd part
IP ⇒ AP	33% (6/18)	40% (10/25)	0% (0/18)	20% (9/46)	60% (15/25)	60% (28/47)
AP ⇒ ∅	17% (7/41)	23% (11/47)	0% (0/42)	0% (0/27)	26% (9/35)	29% (10/34)
Tonal realization						
Hi ⇒ (Hi)	63% (5/8)	73% (32/44)	4% (1/25)	17% (8/48)	81% (17/21)	52% (17/33)

Again speakers 1F and 3M pattern similarly, and show a restructuring of the prosodic phrasing of the text with fewer prosodic groups and a non-realization of certain underlying tones. In both parts, several IP boundaries are reduced to lower-level boundaries (i.e., AP boundaries) at fast rate. Speaker 3M has the largest reduction with a reduction of 60% of his IP boundaries in both parts. Similarly, the number of APs is reduced at fast rate, so that more words are grouped in larger APs. For the two speakers, 17–29% of the AP boundaries are deleted in the two parts of the text. An increase of rate also affects the realization of the underlying tonal pattern of the AP. For the two speakers, 52–81% of the initial high tones (Hi) of the AP are not realized at fast rate.

Speaker 2F, in contrast, shows less variation in the prosodic organization of the text at fast rate. She reduces the number of IPs only in the second part of the text, and to a lesser extent (20%) compared to the other speakers. She does not show any reduction in the number of APs in either part of the text. Initial high tones (Hi) are realized nearly as often in fast rate as in normal rate in the first part (only one Hi not realized) and a little less in the second part (8 Hi not realized out of 48). Thus, for this speaker, an increase of rate has no effect on prosodic organization in the first part of the text, and a very small effect in the second part.

So far, we have assessed the prosodic organization only by its intonational cues. But phrasing is also marked by various degrees of final lengthening. To see whether the durational cues of the prosodic organization are maintained at fast rate, we compared the durational characteristics of syllables in the same prosodic position at fast and normal rates. We expect that, even at fast rate, IP-final syllable should still be longer than AP-final syllables, which in turn should be longer than AP non-final syllables. In Fig. 7, the duration of IP-final syllables, AP-final syllables, and AP non-final syllables are plotted for each speaker at normal and fast rates. AP-initial syllables that have an initial high tone (Hi) are not included. The initial high-toned syllable has been claimed not to have a longer duration than the unaccented syllable in French (Pasdeloup, 1990). In our data, initial high-toned syllables are not always significantly different from lengthened AP-final syllables. Therefore, we excluded the initial high-toned syllables from the

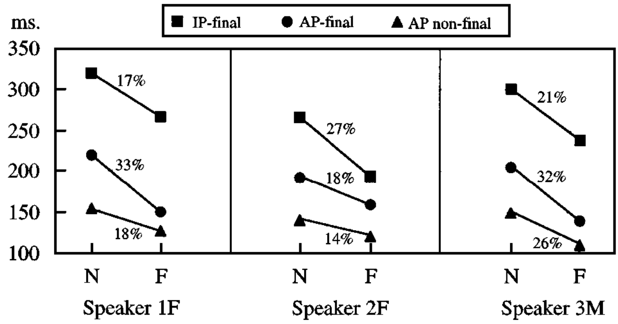

Figure 7. Duration of IP-final, AP-final and AP non-final (also non-initial high tone) syllables at normal (N) and fast (F) rates for the three speakers. For each prosodic position, the percent of reduction of the duration between normal and fast rates is indicated.

comparisons. These durations are not standardized and therefore can be biased by intrinsic and contextual influences of segments. However, since in our data the same syllables are compared in normal and fast rate categories, their intrinsic/contextual properties should contribute in the same way to each category. Results show that the hierarchical organization of syllable duration depending on prosodic position is maintained at fast rate: syllable duration for the three speakers' productions at the three prosodic positions differs significantly at fast rate ($F(2, 1057) = 464.7$, $p < 0.0001$). Thus, at fast rate, IP-final syllables are still significantly longer than AP-final syllables, which in turn are significantly longer than AP non-final syllables (all pairwise comparisons, done with a Scheffé post-hoc test, are significant at $p < 0.0001$ for every speaker).

However, the degree of reduction in the duration is not the same across the three prosodic positions. Speakers 1F and 3M again pattern together by reducing the duration of AP-final syllables (1F: 33%; 3M: 32%) more than that of syllables in other prosodic conditions (IP-final and AP non-final reduction for 1F: 17–18% and 3M: 21–26%). Speaker 2F, on the other hand, shows a degree of reduction that varies linearly with the prosodic hierarchy (i.e., the degree of lengthening): IP-final syllables are reduced the most (27%), AP-final syllables are reduced less (18%), and AP non-final syllables are reduced the least (14%). Thus, for this speaker, the longest syllables (IP-final) are reduced the most, while for the other two speakers, AP-final syllables (medium lengthening) are reduced the most.

4. Discussion

4.1. *Articulation of intonation at fast rate*

In this section we will review results found in both segmental and suprasegmental studies in the literature and compare them with our observations. An increase in rate means a reduction in the time of articulation. As reviewed in the literature on "segmental" articulations, a shortening of articulation time can be achieved in three ways: successive gestures can be realized in a shorter time if the magnitude of their movement is reduced, if the velocity of their movements is increased, and/or if the overlap between gestures is increased. In this paper we have looked at modifications in intonation at fast rate, assuming that an f_0-contour is the output of a laryngeal/subglottal articulation. Here the articulatory gesture is not the spatial displacement of one articulator (such as the tongue) toward another, but rather a variation in the frequency of vibration of the vocal folds. Therefore, movement displacements correspond to the variation between high and low frequencies of vibration; velocity is the speed of the change between high and low frequencies; overlap is the phasing of the gestural activation for high and low frequencies of vibrations.

4.1.1. *Variation in displacements and/or velocity of pitch movements at fast rate*

Segmental studies have shown that velocities and displacements of articulatory movements can be affected by rate increase, and this result varies across studies, articulators, and speakers. For example, Kuehn & Moll (1976) found that one out of five subjects always increased velocity, two usually increased velocity, and two reduced displacements with a concurrent decrease in velocity. Regarding the effect of rate on intonation, we found that an increase of rate is correlated with a *reduction in pitch displacements*. Pitch displacements are generally smaller at fast rate, although this overall reduction is not always significant. This lack of significance can be explained by a larger variation in displacement within the text, at each rate, than that between rates. A reduction of pitch displacement at fast rate has also been found in German (Kohler, 1983) for all three speakers studied, and in Dutch (Caspers, 1994) for one subject (the other subject increased pitch displacements at fast rate).

Movement displacement, duration, and velocity do not vary independently of one another. Several studies have shown that in English there is a strong correlation between displacements and peak velocity, so that larger displacements are generally faster. For example, comparing English to French, Vatikiotis-Bateson & Kelso (1993) showed that the relation between peak velocity and displacements of the lower-lip–jaw complex is less tight in French, and is a poor parameter for distinguishing stress from unstressed gestures. Although it is not obvious that every syllable defined as "stressed" in their study corresponds to a particular pitch event, these results can be compared to the small and inconsistent variations that we found in the average velocity of the pitch movements.

We observed that an increase in rate induces small changes in the average velocity of pitch displacement at fast rate, and that these changes involved either an increase or a decrease in velocity. In the cases of reduction in the average velocity of pitch movement, the reduction in the sharpness of f_0 slopes may result from the reduction in displacement (as found by Kuehn & Moll, 1976). In a few other cases, average velocity of pitch displacements was increased (meaning that the slopes of f_0 movement were sharper) but this increase was never significant. Therefore, it seems that in our study the acceleration in the transitions between successive pitch targets at fast rate was achieved mostly by a reduction in the pitch displacements, rather than by an increase in the velocity of f_0 slopes. Results concerning the velocity of pitch movements are also not consistent across studies and languages. Kohler (1983) found that in German fall–rise f_0 glides are generally leveled at fast rate, meaning that f_0 slopes are less sharp. On the other hand, Caspers (1994) found that, for the accent-lending rise in Dutch, both speakers shortened and steepened the rising pitch movement at fast rate, whereas for the accent-lending fall, the shape of the f_0 contour was preserved at fast rate. An explanation of the difference between languages could be that pitch movements are intrinsically faster in French and German than in Dutch. t'Hart, Collier & Cohen (1990) reported that the most abrupt pitch movement in Dutch intonation is 50 semi-tones per second. In our data, it is in fact the case that the most abrupt pitch movements at normal rate are faster than those in Dutch: 74 st/s for speaker 1F, 89 st/s for speaker 2F, and 88 st/s for speaker 3M. Hence, it may be physically easier for Dutch speakers to make pitch movements steeper when time is short in order to keep them large. Of course, more data are needed to confirm the hypothesis that pitch movements in French are faster than in Dutch.

4.1.2. *Pitch target undershoot at fast rate: linear rescaling, overlap or deletion of pitch gestures*

In Lindblom's (1963, 1964) terms, *target undershoot* is inversely proportional to the duration of the lapse between successive motor commands, allowing or not allowing the articulators to reach their target before the next set of commands arrives. Laryngeal articulation can respond to the same time constraints by undershooting successive high- and low-pitch targets. Ohala and Ewan (1973) and Sundberg (1979) reported that, in singing, f_0 raising takes longer than f_0 lowering. Following Lindblom's reasoning, if a command toward a minimum f_0 target arrives at the muscles before the completion of the preceding f_0-raising movement, undershoot of f_0 maxima is likely to occur. Moreover, if the raising movement is intrinsically longer than the falling movement, this raising movement has more chance to be cut off before the target is reached. This hypothesis of duration-dependent undershoot of high targets is supported by the speech behavior of speakers 1F and 3M, who both reduce their maxima more than their minima. However, this hypothesis does not explain the results for speaker 2F, who undershoots (raises) her f_0 minima rather than her maxima. Kohler (1983) also found raising of f_0 minima at fast rate for his three German speakers. In this study, only four pitch events are taken into account: f_0 maximum in the first word in the sentence, which is also the highest peak in the sentence, and three more points in the sentence which are all valleys. At fast speech rate, all points except the peak are raised, resulting in a reduction of pitch displacements. In sum, it seems that when the rate is faster, either high- or low-pitch targets can be undershot, and which target is undershot seems to be speaker-dependent. Moreover, an increase of rate is also found to result in a general shift in frequency of the whole pitch range. Again the direction of this shift seems to be speaker- and/or language-dependent: we found global lowering of both maxima and minima for speakers 1F and 3M, while Caspers (1994) found global raising of both maxima and minima at fast rate for Dutch.

The reduction of articulatory displacements, leading to the undershooting of pitch gestures, can be driven by two different mechanisms: (1) a linear rescaling, in which the size of the articulatory gestures is scaled in proportion to change in duration, or (2) an overlap of articulatory gestures, in which the reduction of displacements is a function of the degree of overlap between successive (not-rescaled) gestures. For the latter case, the competing demands on one articulator (the vocal folds) made by two opposing gestures (H- and L-gestures) could cause one to be cut off or truncated by the other (Harrington, Fletcher & Roberts, 1995), and could result in the undershoot of one pitch target.

Another possible strategy to increase rate is to simply *delete* some of the gestures because there is no time to articulate them. Munhall and Löfqvist (1992) observed that the two laryngeal gestures corresponding to two voiceless consonants in English can be realized at fast rate with a single smooth laryngeal gesture. Two possible explanations can account for that phenomenon: "… the fastest utterances that exhibit only a single smooth glottal movement could still be composed underlyingly of two separate laryngeal gestures. … On the other hand, at some degree of overlap a reorganization may occur and a single laryngeal movement may be 'programmed' … " (p. 122). In our data, the underlying initial high tone (Hi) of an AP is often not realized at fast rate (at least for two speakers). This raises the question of whether the AP tonal pattern is still composed of four underlying gestures /L H L H/ at fast rate, but the initial Hi is not realized due to complete overlap of this H-gesture with the adjacent L-gestures, or whether the AP pattern is reorganized at fast rate so that AP has only two underlying gestures, L and H. We are inclined to favor the first hypothesis because non-realization of the initial high tone in French is not limited to a fast speech rate condition. In Jun & Fougeron (1995), we observed that the initial high tone is often not realized when the number of syllables within an AP is fewer than four. When the number of syllables is three, a trace of the initial high tone is sometimes seen with a small rise (though it may also appear as a full rise; see Vaissière, 1974; Lucci, 1983; Rossi, 1985). When the number of syllables is one or two, the initial Hi is not realized. This is parallel to Munhall & Löfqvist's (1992) findings on the merger of glottal opening gestures of two voiceless consonants due to time constraints. Traces of those gestures showed cases with two separate peaks, other cases with one large peak with a small shoulder, and instances of a single peak only. If there were reorganization of the tonal pattern at fast rate, we would need to suppose that this reorganization takes place each time the temporal domain of realization of the f_0 contour is shortened (either due to a rate increase or to a small number of syllables). We believe that our finding that some initial high tones remain the same at fast rate suggests that the initial high is present underlyingly at fast rate.

The status of the initial high tone in French, however, needs further study because it is constrained not only by the rhythmic factor and duration as we stated before, but also by other factors such as speech style. For example, the initial rise is more likely to occur in reading and conference styles of speech ("accent didactique"), but less in spontaneous speech (Vaissière, 1971; Lucci, 1983).

4.1.3. *Adding an intonation tier to the task dynamics model of speech production*

The increasing body of data on the influence of prosody on articulatory gestures has suggested the need to introduce a prosodic module in a model of speech production (see e.g., McGowan & Saltzman, 1995; Saltzman, 1995). Our results push this idea forward on the basis of the similarities between the effects of an increase of speech rate in the segmental and suprasegmental domains. We propose that a prosodic tier should be added to a Task Dynamic model of speech production (Saltzman & Munhall, 1989; Browman & Goldstein, 1990; Saltzman, 1995) to model the production of intonation. This tier would be superimposed over the existing tiers since the domain of realization of the "intonation" gestures spans a constellation of gestures. Here we will sketch some of the basic ideas of this proposal, although it is still in a preliminary stage.

In Browman & Goldstein's (1990, 1992) Articulatory Phonology model, the Glottis is one of the articulators considered along with the Velum, the Tongue Body, Tongue Tip, etc. This articulator has one tract variable: glottal aperture, We propose to replace the Glottis articulator by a Vocal Fold articulator. For this Vocal Fold articulator, we have two tract variables: one is the "glottal aperture degree' and is the same as that in the original model (with the three modes given in Browman & Goldstein, 1990; Appendix B, p. 373); the other tract variable is the "vocal fold vibration frequency". For this tract variable, we propose two values: "high" and "low" (however, it is possible that a "mid" value is necessary for languages with a lexical mid tone). Thus, intonation would be modeled by a H-gesture (or "high frequency vibration-gesture") and an L-gesture (or "low frequency vibration-gesture"). Using this model, we can represent the articulation of intonation and its modification observed at fast *vs.* normal rate as schematically shown in Fig. 8. Here we represent the H- and L-gestures as triangles showing the movement toward the H and L targets.

Fig. 8(a) represents the basic pattern of an AP at normal rate, which consists of a succession of an L-gesture, H-gesture, L-gesture, and H-gesture without overlap. Fig. 8(b) represents the undershoot of f_0 maxima and reduction in pitch displacements at fast rate as observed for speakers 1F and 3M. In this representation, the second L-gesture is phased earlier relative to the onset of the first H-gesture. As a consequence, the overlapped H-gesture is truncated. The rising displacement is reduced and the high target is undershot. In this example, it is the initial H-gesture that is truncated, but this can happen to any H-gesture. Similarly, although it is not shown here, an L-gesture can be truncated by adjacent H-gestures (as for speaker 2F). Fig. 8(c) represents the "non-realization" of initial high tone in an AP. This case is a more dramatic version of the overlap presented in Fig. 8(b). The phasing of the second L-gesture is so early relative to the first H-gesture that this H-gesture is totally overlapped. As a result, the H-gesture is hidden by the L-gesture.

By assuming H- and L-gestures we can explain the output patterns and modifications common to laryngeal and supralaryngeal articulators. Further research is needed to develop this model in order to understand the mechanism involved in the adjustments of both the Vocal Fold articulator and other articulator in speech.

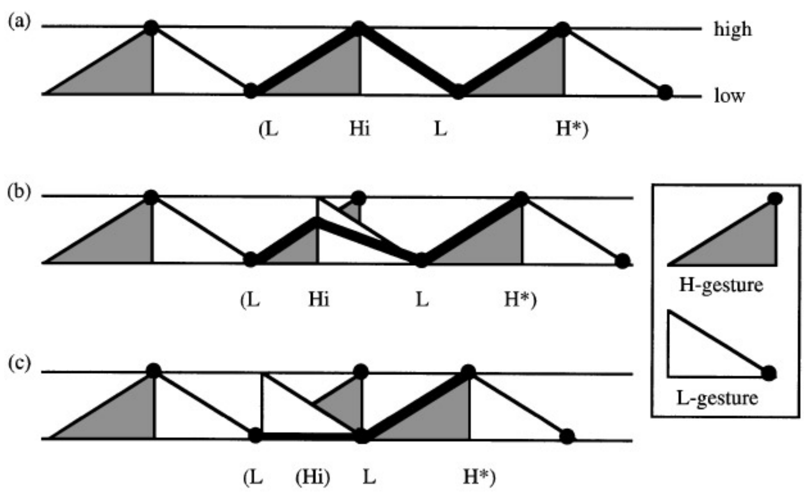

Figure 8. Schematic representations of the model of articulation of intonation, with H-gesture (shaded triangle) and L-gesture (white triangle). Black dots indicate the pitch target for each gesture. The thick black line represents the f_0 contour. In (a), four gestures (L, H, L, H) are successfully realized without overlap; in (b) H-gesture is partially overlapped by the following L-gesture; and in (c) H-gesture is completely overlapped with the following L-gesture, thus realized as [LLH]. The tonal transcription of the AP is shown under each gestural representation.

4.2. *Variation in the compressibility of pitch movement at fast rate*

As observed in segmental studies, the strategies used to increase speech rate are variable. In our data, we have shown that the reduction in pitch range and pitch displacements varies depending on the speaker (two patterns for three speakers) and the position of the speech material in the text (first *vs.* second parts). In this section we will discuss constraints on the compressibility of pitch movements in order to explain these differences.

First, consider the variation in strategy between the first and the second parts of the text for speaker 1F. While the other two speakers keep the same strategy throughout the text, speaker 1F reduced her pitch displacements and pitch range in the first part, but not in the second part at fast rate. This variation can be explained by her use of pitch range at normal rate. If we look at her pitch range at normal rate, it appears that it varies according to position in the text, being wider in the first part (233 Hz) than in the second part (178 Hz). A declination of f_0 from early to late in the speech stream can partially explain this phenomenon: the f_0 top line (plateau) lowers gradually while the f_0 base line is quite stable (see Vaissière, 1983, for a review, but see also Sluijter & Terken, 1993). Therefore, in the second part of the text, the pitch range is already compressed at normal rate. At fast speech rate, it could be the case that pitch displacements and pitch range cannot be further reduced because the pitch range in this part of the text is too narrow to be further compressed. However, it is not the case that pitch displacements are not reduced because the pitch range in the second part is small *per se*. If it were, we should have seen a comparable lack of pitch reduction in the speech of the male speaker who has a smaller pitch range than the female speakers. On the contrary, the male speaker reduced his pitch displacements and pitch range at fast rate just as much as the female speaker did.

Evidence of a "compressibility effect" is also found in the behavior of this male speaker. At fast rate compared to normal rate, the reduction of his pitch range is smaller in the second part of the text than that in the first part. This can be explained by the fact that, at normal rate, his pitch range is already reduced in the second part of the text compared to the first part. Similarly, variation in the degree of compressibility can explain the pattern of reduction of f_0 maxima at fast rate. In Fig. 9, the amount of reduction of f_0 maxima at fast rate is plotted against the height of these maxima at normal rate for the two speakers showing lowering of f_0 maxima at fast rate (speakers 1F and 3M). There is a linear relationship between the reduction of the pitch targets and their f_0 height at normal rate: the higher the maximum, the more it is reduced at fast rate ($r^2 = 0.6$ for the two speakers). In other words, lower pitch maxima are less reduced than higher pitch maxima.

In sum, the production of intonation at fast rate is constrained by what appears to be a saturation effect. Small displacements, for the lower high targets or for the second part of the text for speaker 1F, are not further reduced at fast rate. It is possible that these displacements are small enough to be fully realized, even at fast rate. The resistance to further compression of pitch movement at fast rate could also result from a constraint on the minimal size that a pitch movement must have (or the minimal height a high tone must have) in order to be perceived as a high tone.

Second, consider the difference in strategy used by the speakers in the choice of which pitch target to reduce. We observed two strategies for three subjects: two subjects (1F and 3M) lowered f_0 maxima more than they lowered f_0 minima; one subject (2F) raised f_0 minima and kept f_0 maxima constant. Although only one of our speakers undershot

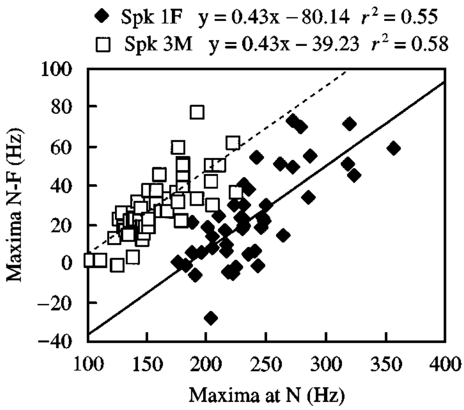

Figure 9. Difference between the height of f_0 maxima at normal and fast rates (maxima N-F in Hz) relative to the height of the maxima at normal rate (maxima at N in Hz).

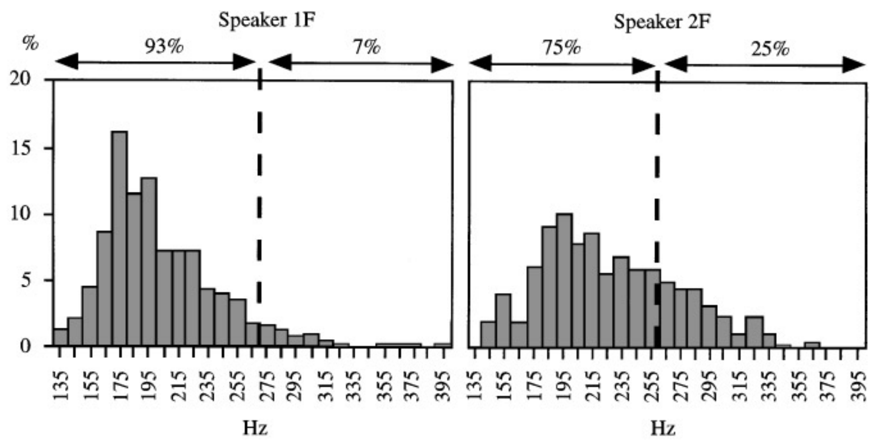

Figure 10. Distribution of the pitch targets for speakers 1F (left panel) and 2F (right panel) depending on their pitch range. The mid-value of their pitch range (mid-range) is shown by the vertical line in each graph. X-axis refers to frequency in Hz and Y-axis refers to the percentage of pitch targets occurring at these frequencies.

pitch minima, this method of reducing pitch displacements was also found in Dutch for one of the two speakers (Caspers, 1994) and in German for all three speakers studied (Kohler, 1983). Thus, we might ask why speakers vary in their choice of which pitch target to reduce at fast rate. The distribution of the pitch targets in the subject's pitch range can offer a plausible explanation for these different strategies. Here we will focus on the two female speakers since they employ different strategies, even though they have a very similar pitch range. Fig. 10 presents the distribution of the pitch targets (both minima and maxima) within the pitch range of speaker 1F (left panel) and speaker 2F (right panel). The vertical dashed line marks the mid-range in each distribution. The pitch targets of speaker 2F are more equally distributed around the mid-range (75% of the pitch targets under, and 25% above) than those of speaker 1F. For speaker 1F, 93% of the pitch targets are concentrated in the lower part of her pitch range. Since speaker 1F uses the lower part of her pitch range more often, the movements toward the high targets located in the upper range may be more extreme, in that they require larger displacement than the movements restricted to the lower region. Thus, when time is short, the targets at the extreme end of the upper range are more likely to be reduced.

A comparable phenomenon has been observed by Kuehn (1976) for velic movements. Looking at velum displacements, he found that his two speakers varied in which movements were reduced at fast rate. One speaker produced less extreme-high positions, keeping low velic positions unaffected, whereas the other speaker produced less extreme-low positions, keeping high velic positions unaffected. Kuehn also noticed that these two speakers were making different use of the range of available velic movement at normal rate. The speaker who reduced the high velic position at fast rate was the one raising his velum beyond the level necessary for velopharyngeal closure at normal rate. In contrast, the other speaker was raising his velum at normal rate just high enough to achieve velopharyngeal closure. Hence, for this speaker, it would not have been possible to reduce further his high velum position at fast rate without compromising the velopharyngeal closure. Comparing this observation with ours, it seems that the movements reduced at fast rate are the movements that are more extreme at normal rate, and which can therefore be compressed when time is short.

Variation in the choice of reducing low or high pitch targets can also be the consequence of prosodic reorganization at fast rate. We observed that speakers 1F and 3M reduced 33–60% of their IP boundaries and 17–29% of AP boundaries. Since most boundaries were high tones, this contributes to the reduction of high targets for these speakers. Along the same lines, maxima may not have been reduced for speaker 2F, since she preserved most of the H tones (AP boundaries and initial high tones) at fast rate.

Finally, interspeaker variation can be explained by an obvious difference in speech styles. Speaker 2F is an especially careful speaker. At normal rate, her articulation rate is fast, but she realizes more pitch events (Hi and IP boundaries) than the other female speaker, for example. At fast rate, she increases her speaking rate to a lesser extent than the other speakers, and shows much less variation both in the shape of the f_0 contour and in the prosodic organization of the text. This speaker also shows very little segmental reduction at fast rate, and produces schwa (at normal and fast rates) more often than most Parisian French speakers do in conversational speech (e.g., in "arriverait", "finalement", "enveloppé", "toutes ces forces"). Thus, the production of this speaker is quite similar at fast and normal rates. On the other hand, speaker 3M varies drastically between the two rate conditions. At normal rate, his prosodic phrasing and tonal realization is as rich as that of speaker 2F. At fast rate, however, this speaker shows the fastest articulation rate and the largest reduction in the shape of the f_0 contour and prosodic organization. Numerous segmental reductions were also observed in this speaker's production at fast rate (e.g., "le plus fort ⇒ l'plus fort", "regardé ⇒ r'gardé", "ils sont ⇒ i sont").

4.3. *Prosodic reorganization at fast rate*

In addition to modification in the phonetic realization of intonation, we found that the prosodic structure of an utterance can be reorganized at fast rate. That is, the prosodic grouping (phrasing) and tonal pattern can be modified when the rate is increased. For example, IP becomes AP, and an AP boundary can be deleted. As a result, a syllable carrying a particular phrasal (H*) or boundary (H%) tones at normal rate may be associated with a different tone at fast rate. An example of this is shown in Fig. 2 with the letter 'B'. In this example, two APs, ("qui arriverait") and ("le premier"), become one AP ("qui arriverait le premier"), and the final vowel [ɛ] in *arriverait*, associated with H* at normal rate, is toneless at fast rate. This means that the number of syllables in one AP is increased and tone-to-syllable associated is reorganized. Changing the prosodic grouping also means reorganization of the information structure: at fast rate more words can group into one prosodic unit, thus fewer prosodic units are observed. This reduction in the number of prosodic units has also been found in French (Vaissière, 1992; Lucci, 1983), Korean (Jun, 1993), and Dutch (Caspers, 1994). Even though the speech material is quite different in these studies, the reduction of AP boundaries in our study, for the speakers reducing these boundaries (19–29%), is similar to that observed in Korean (24%), and the reduction of IP boundaries in our study (20–60%) is similar to that observed in Dutch (56%). However, while the deletion or reduction of boundaries is the result of prosodic reorganization at fast rate, we propose that the absence of initial high tone (Hi) in an AP is not. As explained before (see Section 4.1), it could instead be due to the phonetic realization of the tonal pattern of an AP (with more overlap between H- and L-gestures at fast rate).

Although the prosodic organization is modified at fast rate, it is not the case that the temporal organization of the phrases is disturbed at fast rate. It is well known that, at normal rate, IP-final syllables are longer than AP-final syllables, which in turn are longer than AP non-final syllables (Hirst & Di Cristo, 1984; Pasdeloup, 1990). This hierarchical temporal organization of final lengthening is common cross-linguistically (see Vaissière, 1983, 1992; Wightman, Shattuck-Hufnagel, Ostendorf & Price, 1992, for English). When we compared the syllables that maintained the same prosodic position at fast and normal rates, we found that the hierarchical relation between prosodic units based on lengthening was preserved at fast rate. At fast rate, IP-final syllables are still longer than AP-final syllables, which in turn are still longer than the non-final syllables of the AP.

However, the reduction in the duration of the syllables at fast rate is not similar across prosodic positions. Speakers 1F and 3M reduce the duration of AP-final syllables more than that of IP-final or AP non-final syllables. These speakers are the ones who reduce or delete the most IP boundaries at fast rate. Therefore, we can expect that the few IP boundaries that are maintained at fast rate are important boundaries. Consequently, the duration of the syllables at these boundaries may be less reduced in order to preserve the strength of the boundary. For speaker 1F, we showed earlier that f_0 height at these IP boundaries did not change much at the fast rate (cf. Fig. 5a). Caspers and van Heuen (1991, 1995) also found that boundaries important for discourse structure are less reduced in Dutch. On the contrary, AP boundaries seem to be less important in discourse structure (shown by their high percentage of deletion at fast rate), and their duration is more free to reduce at fast rate (33% for 1F and 32% for 3M). As a consequence, AP-final and AP non-final syllables are less differentiated by duration at fast rate. On the other hand, for speaker 2F, who did not modify the prosodic grouping in the first part of the text and modified it only to a small extent in the second part, the linguistic weight of the boundaries (IP or AP) is not a factor in duration reduction. Rather, the degree of reduction is a function of the duration at normal rate: longer syllables are more reduced at fast rate.

In sum, we have shown that the prosodic grouping of an utterance can be reorganized at fast rate, but is speaker dependent. The pattern of prosodic reorganization is cued by the intonational pattern. At all prosodic levels, we observed a reduction in the number of prosodic groups at fast rate. Also, we showed that even with compression in the temporal domain (rate increase), the durational cues for marking the prosodic structure of an utterance are generally maintained at fast rate.

5. Conclusion

In this paper, we have shown that, in French, an increase in rate affects intonation both in its phonetic realization and prosodic organization. Rate increase induces reduction of pitch range and pitch displacements between maxima and minima pitch targets, with little change in the velocity of pitch movements. The prosodic organization of intonation is modified: some phrase boundaries are reduced or deleted leading to fewer phrases. The realization of the underlying tones is also affected by speech rate. We also showed that these modifications of intonation at fast rate vary depending on the speakers and on the position in the text. This confirms our expectation that intonation of a text provides more variation in pitch range and phrasing than that of an isolated sentence. (The inventory of boundary tones in the text was the same as that in a sentence, i.e., L% or H%). We believe that our study has shown the need to look at both sides of intonation—phonetic realization and phonological structure—when studying intonation. Finally, we showed that at fast rate some dynamic properties of the f_0 contour are modified in a way comparable to the modifications found for other articulations. We have tried to model the observed modifications of intonation using L- and H-pitch gestures. We hope this will lead to further research in developing a model of speech production that includes prosody.

An earlier version of this paper was presented at the XIIIth meeting of the International Congress of Phonetic Sciences, Stockholm, Sweden, in 1995. We wish to thank Pat Keating, Elliot Saltzman, Jacqueline Vaissière, Victoria Anderson, Peggy MacEachern, the three reviewers and the editor, Patrice S. Beddor, for helpful comments and suggestions, and J. Caspers and V. van Heuven for providing their articles. We are also grateful to the speakers for their participation in this experiment. The first author was supported by Allocation de recherche M.R.T. granted to the D.E.A. de Phonétique de Paris.

References

Abbs, J. H. (1973) The influence of the gamma motor system on jaw movements during speech: a theoretical framework and some preliminary observations. *Journal of Speech and Hearing Research*, **16,** 175–200.

Bartkova, K. (1991) Speaking rate modelization in French: application to speech synthesis. In *Proceedings of the XIIth International Congress of Phonetic Sciences*, **3,** 482–485

Beckman, M. & Pierrehumbert, J. (1986) Intonational structure in Japanese and English, *Phonology Yearbook*, **3,** 255–309

Browman, C. P. & Goldstein, L. (1990) Tiers in articulatory phonology, with some implications for causal speech. In *Papers in laboratory phonology I: Between the grammar and physics of speech* (J. Kingston & M. Beckman, editors) pp. 152–178. Cambridge: Cambridge University Press

Browman, C. P. & Goldstein, L. (1992) Articulatory phonology: an overview, *Phonetica*, **49,** 155–180

Brubaker, R. S. (1972) Rate and pause characteristics of oral reading, *Journal of Psycholinguistic Research*, **1**(2), 141–147

Caspers, J. (1994) *Pitch movements under time pressure: effects of speech rate on the melodic marking of accents and boundaries in Dutch.* Doctoral dissertation, Holland Institute of Generative Linguistics, University of Leiden

Caspers, J. & van Heuven, V. J. (1991) Phonetic and linguistic aspects of pitch movements in fast speech in Dutch. In *Proceedings of the XIIth International Congress of Phonetic Sciences*, **5,** 174–177
Caspers, J. & van Heuven, V. J. (1993) Effects of time pressure on the phonetic realization of the Dutch accent-lending pitch rise and fall, *Phonetica*, **50,** 161–171
Caspers, J. & van Heuven, V. J. (1995) Effects of time pressure on the choice of accent-lending and boundary-marking pitch configuration in Dutch. In *Proceedings of Eurospeech*, **2,** 1001–1004
Crystal, T. H. & House, A. S. (1990) Articulatory rate and the duration of syllables and stress groups in connected speech, *Journal of the Acoustical Society of America*, **88**(1), 101–112
Delattre, P. (1966) Les dix intonations de base du français, *The French Review*, **40**(1), 1–14
Engstrand, O. (1988) Articulatory correlates of stress and speaking rate in Swedish VCV utterances, *Journal of the Acoustical Society of America*, **83**(5), 1863–1875
Gay, T. (1981) Mechanisms in the control of speech rate, *Phonetica*, **38,** 148–158
Grosjean, F. & Deschamps, A. (1975) Analyse contrastive des variables temporelles de l'anglais et du français: vitesse de parole et variables composantes, phénomènes d'hésitation, *Phonetica*, **31,** 144–184
Harington, J., Fletcher, J. & Roberts, C. (1995) Coarticulation and the accented/unaccented distinction: Evidence from jaw movement data, *Journal of Phonetics*, **23,** 305–322
Hirschberg, J. & Pierrehumbert, J. (1986) The intonational structuring of discourse. In *Proceedings of the 24th annual meeting of the Association for Computational Linguistics*, New York, NY, 136–144
Hirst, D. & Di Cristo, A. (1984) French intonation: a parametric approach, *Die Neueren Sprachen*, **83**(5), 554–569
Hirst, D. & Di Cristo, A. (in press) *Intonation systems: a survey of twenty languages*. Cambridge: Cambridge University Press
Jun, S.-A. (1993) *The phonetics and phonology of Korean prosody*. Ph.D. dissertation, The Ohio State University
Jun, S.-A. & Fougeron, C. (1995) The accentual phrase and the prosodic structure of French. In *Proceedings of XIIIth International Congress of Phonetic Sciences*, **2,** 722–725
Keller, E. & Zellner, B. (1995) A statistical timing model for French. In *Proceedings of the XIIIth International Congress of Phonetic Sciences*, **3,** 302–305
Kent, R. D. & Moll, K. L. (1972) Cinefluorographic analyses of selected lingual consonants, *Journal of Speech and Hearing Research*, **15,** 453–473
Kohler, K. (1983) F0 in speech timing., *Arbeitsberichte Institut für Phonetik, Universität Kiel* (W. J. Barry & K. J. Kohler, editors), **20,** 57–97
Koopmans-van Beinum, F. J. & van Donzel, M. E. (1996) Discourse structure and its influence on local speech rate. In *Proceedings of the Univ. of Amsterdam*, **20,** 1–11
Krakow, R. (1993) Nonsegmental influences on velum movement patterns: syllables, sentences, stress, and speaking rate. In *Phonetics and phonology: Nasals, nasalization, and the velum* (M. Huffman & R. Krakow, editors) Vol. 5, pp. 87–116. New York: Academic Press
Kuehn, D. P. (1976) A cineradiographic investigation of velar movement variables in two normals, *Cleft Palate Journal*, **13,** 88–103
Keuhn, D. P. & Moll, K. (1976) A cinefluographic investigation of CV and VC articulatory velocities, *Journal of Phonetics*, **3,** 303–320
Liberman, M. (1975) *The Intonational System of English*. Ph.D. dissertation, MIT [Reproduced in 1978 by the Indiana University Linguistics Club, Bloomington]
Lindblom, B. (1963) Spectrographic study of vowel reduction, *Journal of the Acoustical Society of America*, **35**(11), 1773–1781
Lindblom, B. (1964) Articulatory activity in vowels, *STL-QPSR*, **2,** 1–5
Lucci, V. (1983) *Etude phonétique du français contemporain à travers la variation situationelle (début, rythme, accent, intonation, e muet, liaisons, phonèmes)*, Publications de l'Université des langues et lettres de Grenoble.
Malécot, A., Johnson, R. & Kizziar, P.-A. (1972) Syllabic rate and utterance length in French, *Phonetica*, **26,** 235–251
McGowan, R. S. & Saltzman, E. L. (1995) Incorporating aerodynamic and laryngeal components into task dynamics, *Journal of Phonetics*, **23,** 255–269
Mertens, P. (1993) Intonational grouping, boundaries and syntactic structure in French. In *ESCA Workshop on Prosody, Lund Working Paper*, **41,** 155–159
Miller, J. Grosjean, F. & Lomanto, C. (1984) Articulation rate and its variability in spontaneous speech: A reanalysis and some implications, *Phonetica*, **41,** 215–225
Munhall, K. & Löfqvist, A. (1992) Gestural aggregation in speech: Laryngeal gestures, *Journal of Phonetics*, **20,** 111–126.
Ohala, J. & Ewan, W. G. (1973) Speed of pitch change, *Journal of the Acoustical Society of America*, **53**(1), p. 345 (Abstract)
Ostry, D. & Munhall, K. G. (1985) Control of rate and duration of speech movements, *Journal of the Acoustical Society of America*, **77**(2), 640–648
Pasdeloup, V. (1990) *Modèle de règles rhythmiques du français appliquées à la synthèse de la parole*. Doctorat Université de Provence, Aix en Provence

Pierrehumbert, J. (1980) *The Phonology and phonetics of English intonation*, Ph.D. dissertation, MIT
Rossi, M. (1985) L'intonation et l'organisation de l'énoncé, *Phonetica*, **42,** 135–153
Saltzman, E. (1995) Intergestural timing in speech production: data and modeling. In *Proceedings of the XIIIth International Congress of Phonetic Sciences*, **2,** 84–91
Saltzman, E. & Munhall, K. G. (1989) A dynamic approach to gestural patterning in speech production, *Ecological Psychology*, **1,** 333–382
Sluijter, A. M. & Terken, J. M. (1993) Beyond sentence prosody: paragraph intonation in Dutch, *Phonetica*, **50,** 180–188
Sundbeg, J. (1979) Maximum speed of pitch changes in singers and untrained subjects, *Journal of Phonetics*, **7,** 71–79
t' Hart, J., Collier, R. & Cohen A. (1990) *A perceptual study of intonation*. Cambridge: Cambridge University Press
Vaissière, J. (1971) *Contribution à la synthèse par règle du français*. Thèse de troisième cycle, Grenoble, France
Vaissiere, J. (1974) On French prosody, *MIT RLE Quarterly Progress Report*. No. 114, 21–223
Vaisseire, J. (1983) Language-independent prosodic features. In *Prosody*: *Models and measurements* (A. Cutler & R. Ladd, editors) pp. 53–66 Berlin: Springer
Vaissière, J. (1992) Rhythm, accentuation and final lengthening in French. In *Music*, *language*, *speech and brain* (J. Sundberg, L. Nord, & R. Carlson, editors). Wenner-Gren, International Symposium Series, **59,** 108–120, Stockholm
Vatikiotis-Bateson, E. & Kelso, J. A. S. (1993) Rhythm type and articulatory dynamics in English, French and Japanese, *Journal of Phonetics*, **21,** 231–265
Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M. & Price, P. (1992) Segmental duration in the vicinity of prosodic phrase boundaries, *Journal of the Acoustical Society of America*, **91,** 1707–1717