PROPERTIES OF FRENCH INTONATION AT FAST SPEECH RATE

Cécile Fougeron and Sun-Ah Jun°**
** Inst de Phon., CNRS URA-1027, Paris III, °* Dept. of Linguistics, UCLA, USA*

ABSTRACT

Effects of fast speech rate on French intonation in text reading are reported. Modifications at fast rate were found in both the shape of F0 curves and the prosodic organization of the text (phrasing and realization of underlying tones), with different patterns across speakers. The results suggest that F0 realization at fast rate is not a simple "speeding up" compared to normal rate but involves a prosodic reorganization.

INTRODUCTION

The realization of intonation patterns has been found to be influenced by several sources of variation, such as pitch range, sentence structure and length [5], or focus. However, very little is known about the effect of speech rate on intonation other than its effect on the number of intonational phrases [2, 8]. As far as we know, no systematic description has been done regarding the phonetic properties and phonological patterns of intonation when its physical domain of realization is shortened by an increase in rate.

So far, the observation of the effects of speaking rate has mostly been restricted to the segmental domain, showing modifications in the temporal, spectral, and articulatory organization of speech. Articulatory data have shown that there are different strategies for overcoming alteration in articulation time; mainly, target undershoot with reduction of the magnitude of movements, and/or increase in stiffness for faster transitions between targets [6, 1]. If we assume that a tune is composed of a sequence of underlying tonal targets [7], we can use the same analysis techniques as used in studies of segmental articulation to observe the effect of fast rate on intonation.

By examining F0 curves, we can test the hypothesis that there is a modification in the physical realization (F0 shape) of the underlying intonation structure as well as a reorganization of the prosodic structure.

METHOD

The text "La bise et le soleil" ("The wind and the sun", given in Table 1) was read by one male and two female Parisian French speakers at self selected normal and fast rates, with three repetitions at each rate. The full text was analyzed for one of the female speakers (1F). For the two other speakers (2F, 3M) the comparison was limited to the first half of the text where most variation between fast and normal rate was found for Speaker 1F.

Acoustic measurements were taken for each tonal pattern at the rising onset and falling offset (hereafter called F0 minima), and the peak (F0 maxima). Duration of accented syllables and intonational phrases were measured.

Qualitative comparison of the prosodic structure is based on the model of French intonation developed in Jun & Fougeron [3] where the lowest tonally defined prosodic level is the Accentual Phrase (/LHLH/), and the highest prosodic level is the Intonational Phrase.

RESULTS AND DISCUSSION

A. Differences depending on the position in the text.

The reduction of duration at fast rate for Subject 1F was about 33% compared to her normal rate. While this reduction is constant for the whole text, the rate acceleration seems to have different effects on F0 depending on the position in the text. In the first part of the text (see Table 1), rate acceleration induced reductions in the shape of F0 curve, as well as some changes in the prosodic organization of the phrases, while in the second part of the text we observed only phrasing differences.

In the first part (Fig. 1), F0 range is reduced at fast rate (32%) with the highest F0 value being lowered and the base line (lowest F0 value) being stable.

In addition, the displacements between L and H targets are also reduced by significantly ($p<.01$) lowering the maxima (17%) more than the minima (9%). Despite this overall lowering of the peaks, the peaks at major prosodic boundaries (#3, 10, 12, 19 in Fig. 1), critical points of prosodic organization, are not reduced. Comparison of F0 movement velocity (displacement Hz/time) shows very little difference between the two rates. The movements are not stiffer at fast rate, with the reduction of displacement being proportional to the reduction of duration.

In the second part of the text (Fig.2), rate acceleration induces a small reduction of F0 range (6%) but no change in displacement with small and equal degree of lowering of maxima (5.8%) and minima (5.5%). The velocity of F0 movements are not affected.

Considering the prosodic organization of the text, we found there is a change in the realization of tonal patterns at fast rate: in an Accentual Phrase /LHLH/ [3], the initial underlying high tone (initial accent) is often not realized (63% and 73% for 1st and 2nd part). We also found fewer prosodic groups at fast rate. Intonational Phrase boundaries were reduced to lower level boundaries (50% and 42% for the first and second part), and some of the lowest boundaries (AP) were deleted (22% for both the first and second part).

In sum, different strategies are adopted in the first and second parts of the text. When comparing the range of variation of F0 targets in the two parts, evidence is found for a saturation effect. In the first part, the F0 range is wide, which allows some variation in the displacement towards F0 targets without changing the prominence relations between successive targets. In the second part of the text, range of variation is reduced with a smaller F0 range (124Hz compared to 213Hz in the first part). A second indication of saturation is found in the fact that the lowering of the maxima at fast rate is proportional to their height at normal rate: as shown in Fig. 3, in both the first (empty square) and the second (circle) part of the text, the higher the peak is, the bigger the reduction is at fast rate (R^2=0.28 at both positions).

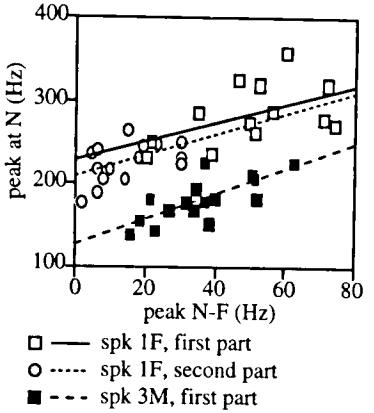

Figure 3: Lowering of F0 peaks between normal (N) and fast (F) rate depending of the height of the peaks at normal rate.

B. Differences between speakers

The three speakers show two patterns both in the way of reducing displacement in F0 movement and in prosodic organization.

The male speaker (3M) shows a pattern similar to Speaker 1F (presented above). With a reduction of 44% from the normal rate, this speaker significantly ($p<.01$) lowers his high and low targets. The lowering of the maxima is correlated with the height of the peak at normal rate (Fig. 3, R^2=0.56). Reduction of displacement is achieved by lowering the maxima (17%) more than by lowering the minima (5%). This speaker also shows a similar reduction in the number of phrases (55% IP and 28% AP reduction) and tonal realization (81% of the initial high tones are not realized).

Speaker 2F shows a totally different strategy. With a rate acceleration of 25% between her fast and normal rate, she doesn't change F0 maxima (2%, $p=0.3$) but significantly raises F0 minima (5%, $p<.01$), compared to the normal rate, as shown in Fig. 4. There is also no difference in velocity or pitch range. This speaker also differs from the others in that she doesn't modify the prosodic organization of the text. At

fast rate, the phrasing is exactly the same as in the normal rate, and the tones are realized as they are in the normal rate.

Thus, at fast rate, this speaker obtains displacement reduction by undershooting her low targets, keeping the high constant, while the other speakers reduce their displacement by lowering high targets.

When we compare the distribution of F0 targets within the range of the speakers, we can observe that for speaker 2F, the targets are equally distributed around the mid-range. For speaker 1F, the dispersion is more concentrated (74% of the L and H targets) in the lower part of her range. Since she uses the lower part of her pitch range more often, movements toward high targets located in the upper range require more energy than the displacement within the restricted lower region. Thus, when the time is limited, at fast rate, the targets at the extreme end of the upper range would be more likely to be reduced (cf. [4]). For speaker 2F, the interpretation is less clear; since she uses both parts of her pitch range equally, her target distribution doesn't provide any indication of a preference for reducing high or low targets. Therefore, there must be other factors to account for her raising of the low targets. Since high targets (mostly boundary tones) are linguistically more important in the prosodic grouping of our text, we can expect she may prefer to modify low targets. This factor may not be the main constraint for speaker 1F whose reduction of the high targets does not jeopardize their prosodic salience.

It should be also noted that the occurrence of target lowering is not determined by the extent of a speaker's pitch range. For example, although the male speaker 3M has a rather small pitch range (139Hz), he shows a high degree of displacement reduction when he accelerates his rate. In contrast, speaker 1F shows little reduction in the second part of the text, where her pitch range is reduced to a value (124Hz) similar to the male speaker's. Thus, the saturation effect we found for speaker 1F must be due to the reduction of her full pitch range at the end of the text, but not because of a small range itself.

CONCLUSION

In this paper, we showed that an acceleration of rate induces a reduction of pitch range and displacement between H and L targets, but no change in the velocity of F0 movements. Moreover, at fast rate, lower prosodic units are regrouped into higher prosodic units leading to fewer phrases, and some of the underlying tones are not realized. We also found that the effects of an acceleration of rate vary across speakers, and depend on the nature of the target tone (H or L), its position relative to the pitch range, its linguistic function (strength of boundary), its position in the text, and the distance between target tones.

ACKNOWLEDGEMENT

The first author is supported by Allocation de recherche M.R.T. granted to the D.E.A. de Phonétique de Paris.

REFERENCES

[1] Gay T. (1981). "Mechanisms in the control of speech rate", *Phonetica* 38: 148-158.

[2] Jun, S.-A. (1993) *The Phonetics and Phonology of Korean Prosody*. PhD. diss. The Ohio State Univ.

[3] Jun S.-A. & Fougeron C. (1995). "The Accentual Phrase and French Prosodic Structure". (this conference)

[4] Kuehn D.P. & Moll K. (1976). "A cinefluographic investigation of CV and VC articulatory velocities", *J. of Phonetics* 3:303-320

[5] Liberman M. & Pierrehumbert J. (1984). "Intonational invariance under change in pitch range and length", in M. Aronoff and R. Oehrle (eds.) *Language Sound Structure*, MIT Press.

[6] Lindblom B. (1963). "Spectrographic study of vowel reduction", *JASA* 35 (11): 1773-1781.

[7] Pierrehumbert J. (1980). *The Phonology and Phonetics of English Intonation*. PhD. diss. MIT.

[8] Vaissière J. (1986). "Variance and invariance at the word level" in Perkell J. & Klatt D. (eds.) *Invariance and Variability in Speech Processes*, L. Erlbaum Assoc., 534-539.

Table 1: Text "La bise et le soleil" and boundary codes used in figures 1, 2, and 4.
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).

Figure 1: spk1F, first part of the text

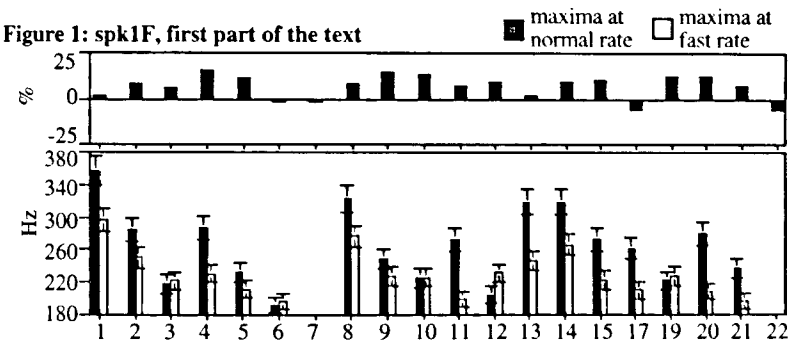

Figure 2: spk1F, second part of the text

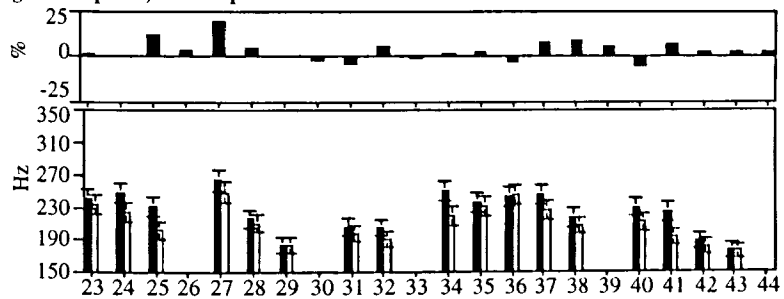

Figure 4: spk2F, first part of the text

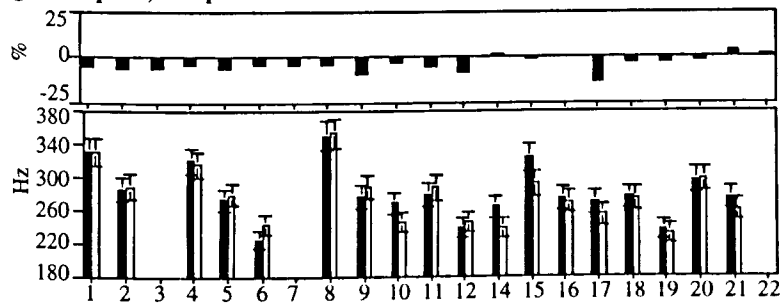

Figure 1, 2, and 4: In each graph, the lower part shows the difference (in Hz.) between normal (black) and fast (white) rate for F0 maxima. The upper part shows the difference (in %) of normal rate (N-F/N) for F0 minima (a positive value = a lowered minima at fast rate).