EFFECTS OF ARTICULATION RATE ON DURATION IN READ FRENCH SPEECH

Danielle DUEZ
Laboratoire Parole et Langage, CNRS, ESA 6057, 13621 Aix en Provence, France
email/duez@lpl.univ-aix.fr

ABSTRACT
The present analysis examined the interactive effects of articulation rate and position in phrases and utterances on the duration of syllables, consonants, and vowels in read French speech. Three major tendencies emerged: 1) Rate changes primarily affected vocalic segments. 2) Non-prominent (penultimate) syllables, consonants and vowels were more stable than their prominent (final) counterparts. This finding, which differs from that obtained for lexical-stress languages, is consistent with the claim that French is not a stress language but a boundary language. 3) At the normal rate, there was no significant difference in mean final-syllable, consonant and vowel duration in the utterance-internal and utterance-final positions. This finding suggests that utterance-final lengthening varies with articulation rate, speech style and speaker, and points out the need for a more comprehensive study of lengthening effects.

1. INTRODUCTION
The influence of articulation rate (i.e. the amount of time spent articulating) whether alone or in conjunction with other factors, on speech segment duration has been investigated in studies whose main aims were to quantify the effect of articulation rate on duration and shed some light on the compression process.

The analysis of the effect of articulation rate on various speech segments has demonstrated that vowels and consonants behave differently for a given change in rate, with vowel duration changing more than consonant duration. This finding has been interpreted as an indication of the constraints on achieving a particular targeted articulatory configuration, stricter in consonant production than in vowel production. Articulation effects have also been shown to be language specific and to depend on the phonological structure of languages. Therefore, our first aim was to determine some temporal features of French by examining the effect of articulation rate on CV syllables, vowels, and consonants in read sentences.

The interaction of articulation rate with the language rhythm adds another level of complexity. In lexical-stress languages, the articulatory gestures of stressed syllables, consonants, and vowels are maintained at the expense of the gestures of unstressed syllables, consonants, and vowels in order to preserve information crucial to lexical access. In French where there is no lexical stress, the rhythmic pattern of a read utterance mainly relies on the prominence given to final syllables at the edge of a prosodic word [1] or breath group [2]. The realization of final prominences is tightly linked to the syntactic structure of the utterance but is also influenced by factors such as constituent length and speech style. The second objective of this experiment was thus to examine how articulation-rate changes interact with phrase-final lengthening. The effects of articulation-rate variations on penultimate and final syllables, vowels, and consonants were examined, and compared in order to determine which syllables and segments (penultimate or phrase-final) are more sensitive to articulation-rate effects.

Final lengthening has been a topic in a wide number of studies which, among other effects, have looked at the magnitude of lengthening in various languages. Final lengthening has been found in word-final and phrase-final positions, as well as at the end of utterances. A tendency toward less lengthening in the word-final and phrase-final positions than in the utterance-final position has been observed in many languages. However, this tendency was questioned in a study on duration in English paragraphs[3] where there was no significant differences between phrase-final syllables and sentence-final syllables, suggesting that utterance-final lengthening in English may surface in sentences produced in isolation. Utterance-final lengthening is a complex phenomenon whose magnitude may be specific to the language, speech-style and speaker. Therefore, our third goal here was to investigate the interacting effects of articulation rate and position in the utterance (utterance-internal and utterance-final) on the duration of penultimate and final syllables, consonants, and vowels in French.

2. EXPERIMENTAL PROCEDURE
2.1. MATERIALS
The speech material consisted of 30 sentences read five times by five French speakers. Each sentence contained one of the 30 trisyllabic prosodic key words, each word containing in turn the target syllable. The target syllable was of the CV type, this syllable structure being prevailing in French [4]. The consonant was one of the following (/p, t, k, b, d, g, l, n, m, f, s, er, v, z, q/) and the vowel was an /a/. The subsequent consonant was a voiceless stop. There were two sets of prosodic key words. In the first set, the target syllable was phrase-final, in the second set, it was the penultimate. The key words were inserted in two carrier sentences with the same number of syllables (10). In the first series of sentences, the key words were placed at the edge of an...
intonation phrase and in the middle of the sentence, and in the second series, they were sentence-final.

2.2. SUBJECTS
Five adult native-French speakers, two men and three women in the age range of 20 to 40 years, were the subjects. None of them had a strong regional accent and none reported any speech or hearing abnormalities. There were ten sessions. The sentences were displayed on index cards. In the first part of the session, speakers were asked to read the sentence at their normal speaking rate. In the second part, they were instructed to read at a fast rate, though intelligible, and to avoid inserting pauses or placing initial prominences on the test tokens. The sentences were recorded in a special recording room at the Parole et Langage laboratory of Aix-en-Provence, France. The recordings were made on a Sony cassette recorder using a microphone. A total of 3000 sentences were obtained (15 consonants x 2 locations within a sentence x two locations within a phrase x 2 speech rates x five speakers x five repetitions).

2.3. ACOUSTIC ANALYSIS
The sentences were digitized at a sampling rate of 16 KHz on a Sun computer. Measurements were carried out on the spectrograms and the oscillograms displayed on the screen, and by listening to selected segments of the waveform in the zone of interest. Test-syllable duration, consonant and vowel duration in the test syllable, and the total duration of each test word and sentence were measured. The compression percentage was calculated by dividing the difference between the normal-speech segment duration and the fast-speech segment duration, by the normal-speech segment duration. This was aimed at comparing the simple and compound effects of rate, phrase location, and utterance on the duration of syllables, consonants, and vowels. The same method was used to calculate the degree of lengthening of final syllables, consonants and vowels in the two rate conditions.

3. RESULTS
3.1. SENTENCE DURATION
As expected, the mean duration at the fast rate (mean=1214 ms, SD=127) was significantly shorter than at the normal rate (mean=1691 ms, SD=221). Similarly, each of the five individual subjects had shorter sentence durations at the fast rate than at the normal rate: normal-rate sentence durations ranged from 1594 ms (Spk2) to 1897 ms (Spk3), and fast-rate sentence durations ranged from 1088 ms (Spk2) to 1343 ms (Spk3). There was more variation across speakers at the normal tempo than at the fast tempo (coefficient of variation: normal speech 0.13; fast speech 0.10). The mean sentence durations obtained for the five speakers were input into a Speaker x Rate ANOVA. The main effects of rate [F(1, 2990)=8018] and Speaker [F(4, 2990)=387] were significant (p=0.0001). The two-way interaction was also significant [F(4, 2990)=20, p=0.0001]), which indicated that speakers did not change speaking rate in the same way.

3.2 RATE, SYLLABLES, CONSONANTS, AND VOWELS.
The duration of syllables, vowels, and consonants varied substantially across rates for all speakers. An ANOVA with Rate (2) and Speakers (5) as factors yielded reliable effects of Rate (syllables, [F(1, 2990)=920], consonants [F(1,2990)=344], and vowels [F(1,2990)=826], p=0.0001) and Speakers [syllables F(4,2990)=56; consonants: F(4,2990)=24; and vowels: F(4,2990)=82; p=0.0001], as well as a reliable interaction between the two factors for syllables [F(4, 2990)=3, p=0.009]. The simple effect of rate proved to be significant for every speaker.

The syllable percentages obtained for the five speakers ranged from 16% to 25%, and the mean was 22%. For consonants, the range was 18% to 24%, and the mean was 21%. For vowels, the percentages ranged from 22% to 34% and the mean was 24%. Higher vowel values indicate a greater likelihood of vowels to undergo rate changes than consonants, which confirms the results reported in the literature. The five speakers had different compression strategies. For example, Speakers 2, 3, 4 and 5 had higher syllable-compression percentages than vowel- and consonant-compression percentages; for Speaker 1, it was the other way around. For Spk 2 and 3 the compression percentages obtained for syllables, consonants and vowels were very close; for Spk4, there was a large difference between the consonant and the vowel compression percentages. A similar variability across speakers with changes greater for vowels than for consonants was reported for French [5] However, the percentages obtained here are higher.

3.3 RATE AND LOCATION WITHIN A PHRASE
The simple effect of rate on duration proved to be significant for syllables [F(1,2980)=2134], consonants [F(1,2980)=448], and vowels [F(1,2980)=1594; p=0.0001]. Similarly, the effect of speaker was found to be significant for the three categories of segments [syllables F(4,2980)=43; consonants F(4,2980)=31.9; and vowels F(4,2980)=159; p=0.0001] as was the effect of position-in-phrase [syllables F(1,2980)=3558; consonants F(1,2980)=821; and vowels F(1,2980)=2440; p=0.0001].

There were also significant interactions between rate and position-in-phrase for syllables [F(1, 2980)=225, p=0.0001], consonants [F(1, 2980)=33, p=0.0001] and vowels [F(1, 2980)=209, p=0.0001], speaker and position-in-utterance for syllables [F(4, 2980)=40, p=0.0001], consonants [F(4, 2980)=12, p=0.0001], and vowels [F(4, 2980)=26, p=0.0001]; speaker and rate for syllables [F(4, 2980)=7, p=0.0001], and vowels [F(4, 2980)=4, p=0.001]; and rate, speaker, and position-in-utterance for syllables [F(4, 2980)=6, p=0.0001] and vowels [F(4, 2980)=7, p=0.0001]. The duration tended to decrease more as a function of rate for final segments
than for penultimate segments, for Speakers 2, 3 and 5 than for Speakers 1 and 4 in the final position, and for Speakers 1, 2 and 5 in the penultimate position.

The compression percentages clearly indicated differences across speakers and position-in-phrase. Changes in duration were smaller for penultimate syllables, consonants and vowels than for final ones: Final syllables, consonants, and vowels declined substantially while penultimate ones tended to remain more stable. Similarly, differences across speakers were greater in the final than in the penultimate position. In the final position, the percentages ranged from 29% to 21% for syllables, 29% to 18% for consonants, and 29% to 23% for vowels. In the penultimate position, the ranges were 20% to 16% for syllables, 19% to 17% for consonants, and 23% to 15% for vowels.

The degree of lengthening of final syllables, consonants and vowels (compared to penultimate ones) did not stay constant across speech rates, for the group and for the individual speakers. Mean-lengthening percentages were smaller for the fast rate (syllables 24.53%; consonant 24.36% and vowels 24.37%) than for the normal rate (syllables 30.74%; consonants 28.55%, and vowels 32.29%). The individual percentages varied substantially across rates for syllables (normal: 22.58% to 36.95%; fast: 20.20% to 28.09%), consonants (normal: 20.88% to 34.62%; fast: 15.45% to 29.88%), and vowels (normal: 25.45% to 38.54%; fast: 19.38% to 28.79%), suggesting different lengthening strategies. For instance, Spk1 reduced consonant duration more than vowel duration at the fast rate while for Spk4 it was just the other way around. The mean lengthening percentages for syllables were lower than those reported for French CV syllables [6] in read texts (around 0.70%). However the lowest lengthening degree found here (20%) corresponds to the perceptual lengthening threshold [7].

3.4. RATE, LOCATION WITHIN A PHRASE, AND LOCATION WITHIN AN UTTERANCE

All four main factors proved to be significant for syllables, consonants and vowels. That is, all durations were significantly shorter at the fast rate than at the normal speaking rate, and in the penultimate position than in the phrase-final position. The durations across speakers and the two utterance positions were also significantly different. The analysis of the $F$-values yielded the following hierarchy in descending order: Position-in-phrase (4719), Articulation rate (2846), Speaker (174) and Position-in-utterance (143). The $F$-values were lower for consonants than syllables or vowels, confirming the fact that consonants are less sensitive to rate and position in phrase and utterance, and are less speaker-specific than syllables and vowels. All of the interactions were significant for syllables and vowels. For consonants, there were no reliable interactions between speakers and position-in-utterance, or any of the other factors. To summarize the most important interactions, segment duration decreased less in the utterance-final position than in the utterance-internal position, and in the penultimate position than in the phrase-final one. Differences in the durations of utterance-final segments and utterance-internal segments also seemed to be rate- and speaker-dependent. At the fast rate, all mean durations of syllables, consonants, and vowels in utterance-final position were longer than the durations of these units in utterance-internal position. At the normal rate, utterance-final syllables and consonants were longer than utterance-internal ones while, the opposite was true for vowels. This pattern, which is not fully consistent with the general tendency reported in the literature, reflects speaker specificity. While Speakers 1, 2 and 4 had longer utterance-final syllables and vowels than utterance-internal ones, the inverse pattern was observed for Speakers 3 and 5 who had longer utterance-internal syllables and vowels than utterance-final ones, the difference being particularly sharp for Spk 3.

The percentages obtained for final syllables, consonants, and vowels were higher than for their penultimate counterparts, confirming the fact that rate-change effects were stronger for the former than for the latter. The mean percentages obtained for final as well as penultimate utterance-internal segments (final: syllables 33%, consonants 27% and vowels: 37%; penultimate: syllables: 20%, consonants 18%, and vowels 22%) were higher than those obtained for final and penultimate utterance-final segments (final: syllables 17%, consonants 18%, and vowels 16%; penultimate: syllables 16%, consonants 17%, and vowels 15%). This is quite consistent with the above finding that syllables, consonants, and vowels underwent greater compression in utterance-internal position than in absolute-final position.

Final lengthening of syllables, consonants, and vowels occurred at the normal and fast rates in the utterance-final position and in the utterance-internal position. Again, the degree of lengthening appeared to be more stable for syllables, consonants and vowels in the utterance-final position than in the utterance-internal position, at both rates. For example, in the absolute-final position, the syllable-lengthening percentages ranged from 44% (Spk1) to 62% (Spk2) at the normal rate and from 42% (Spk4) to 62% (Spk2) at the fast rate. In the internal-utterance position, the ranges were 15% (Spk1) to 63% (Spk3) for normal speech, and 6% (Spk1) to 32% (Spk3) for fast speech. Lengthening percentages under 20% are below the duration perception threshold, which suggests that there was no acoustic realization of the syntactic boundary. Subject percentages across rate conditions confirmed the greater variability of final lengthening in the utterance-internal position. Final-lengthening in the absolute-final position and the utterance-internal position also appeared to be realized in different ways by the five speakers. For speakers 3 and 5 the degree of lengthening was greater on internal-utterance syllables,
consonants, and vowels than on final-utterance syllables, consonants, and vowels, for Speakers 1, 2, and 4, it was the other way around.

4. GENERAL DISCUSSION

The acoustic data obtained here revealed that the observed rate changes primarily involved vocalic segments. Consonantal segments were affected as well, but to a lesser degree. This finding goes against the results reported for a corpus containing the readings of one speaker where consonants and vowels were found to reduce in the same way[8]. In contrast, they are quite in line with the results obtained for texts read at different tempi by six French speakers [1, 5]. More generally, they corroborate those previously obtained for other languages.

The overall duration of sentences in normal and fast speech was shown to vary significantly across speakers. This finding supports the claim that articulation rate is speaker-specific [9]. Moreover, vowels and consonants were not altered in the same way by the five speakers, reflecting different articulatory strategies for achieving faster rates of speech. Speakers seem to have considerable freedom to control the individual components of speech in response to rate changes.

Non-prominent (penultimate) and prominent (final) syllables, consonants and vowels were not equally affected by rate variations. Thus, non-prominence may be a baseline and as such, may be a more stable feature than prominence. This result differs from that obtained for lexical-stress languages and is in line with the claim that French is not a stress language but a boundary language[2].

There was a high degree of sensitivity to articulation rate variations on final syllables. The duration of final syllables consonants, and vowels decreased as articulation rate increased. The extent of the lengthening between penultimate syllables and segments to final ones decreased with rate. Even more, in some cases, the strength of the articulation rate-effects completely eliminated final lengthening. By contrast, penultimate syllables, consonants and vowels had about of the same duration in the two utterance positions, at both rates. This finding confirms the stability of penultimate syllables and segments, which seems to be constant across utterances. At the normal rate, there was no significant mean duration difference on final syllables, consonants, and vowels in utterance-internal or utterance-final position, contrary to the fast rate. The lack of a significant finding for the two positions-in-utterance in normal rate contrasts with the greater degree of lengthening reported for spontaneous speech and read sentences in French [2]. However, it is fully consistent with other results obtained for read French. The final contour was not found to be necessarily longer than the other contours in the sentence and thus could not be differentiated by means of this feature [10]. Accented syllables were 26% shorter in the final position than in the internal position [11]. The same was true for 36% of the accented syllables in the absolute final position; moreover, in some cases, final syllables were shorter than non-final syllables [12].

A closer examination of the present results suggests that the lack of greater lengthening in the final position could be linked to individual strategies. Three of the speakers exhibited greater utterance-final lengthening than utterance-internal lengthening; for the two other speakers, it was just the opposite. These speaker-linked differences are consistent with previous studies reporting a high degree of interspeaker variability in the magnitude and implementation of utterance-final lengthening [1]. Generally, utterance-final lengthening is regarded as a natural phenomenon that reflects the deceleration accompanying the end of motor activity (13). The specificity of some of the results obtained here suggest that utterance-final lengthening varies with articulation rate, speech style and speaker, and points out the need for a more comprehensive study of lengthening effects.

REFERENCES