Some micro-effects of tempo change on Timing in French

Janet Fletcher

The Centre for Speech Technology Research, University of Edinburgh

Abstract

A number of factors influence the temporal organization of connected speech. One such factor, tempo change, is an important variable whose effects on segment duration may highlight important language-specific and language universal features of speech timing. When French speakers are asked to read a passage at increasing tempo from slow to normal to fast, they modify their pause behaviour substantially. In addition, consonant and vowel segment durations are compressed, but not to the same degree. Whilst the overall levels of temporal compression are lower for vowels than the levels of compression reported for other languages, they are still somewhat higher than previously reported for French. A probable reason for the former is the smaller degree of vowel reduction that occurs in French under conditions of increased tempo. This factor along with other compression effects such as syllable, consonant and vowel elision are considered in the light of traditional timing typologies./fR

Introduction

The durational structure of a stretch of connected speech is determined by a number of factors. Speech tempo is one such factor that has received some attention in the experimental phonetics literature. However like the majority of instrumental studies of speech in general, past research has concentrated on a limited sample of languages, most notably English. In addition, the type of corpora analysed have usually been restricted to short controlled utterances. Few researchers have looked at the effects of tempo change over a large stretch of connected speech. Consequently, this study constituted an attempt to examine in detail the effects of tempo change on one feature of timing organization, namely, segment duration in a long stretch of connected speech. These effects can be described as the micro-effects of tempo change. Furthermore, the language under investigation was French. We wished to examine any possible strategies at the segmental level to modify tempo in spoken French, that differed from those reported for other languages.

Micro-effects of tempo change

Previous studies of tempo change have found that the greatest quantitive effect of increasing or reducing tempo is to do with reducing or increasing the amount of pause time relative to articulation time. Smith (1976) and Duez(1983) found this to be the case in French. However there was also evidence of changes in articualtion rate between tempo that can be explained by temporal changes at a micro-level.

A number of authors have investigated the effects of tempo change on the acoustic duration of segments. However, most of these studies have concentrated on corpora composed of isolated sentences and not of larger tracts of connected speech. Whilst the first kind of study provides us with some idea of the effects tempo change on certain categories of acoustic segments, we need to extend these findings to a broader range of speech situations. We are aware of only one controlled study of micro-effects of tempo variation in a passage of spoken French(Smith 1976) Consequently we will report on the findings of some studies of tempo change that have used single utterance-based corpora, in addition to those studies based on larger stretches of read English, Swedish and Dutch.

Gay(1978) and Port(1984) for American English, Lindblom(1963) for Swedish, and Nooteboom and Shs(1972) for Dutch, found that overall decreases in syllable duration during rapidly delivered speech were reflected primarily in the durations of the constituent vowels, and only secondarily in the durations of the constituent consonants. Port(1984) for example, provided an explanation for this based on universal mechano- inertia effects of increasingly rapid articulation. Consonant articulations are generally more complex than vowel articulations, in so far as they imply the attainment of a specific constriction, or closure at some point, or points in the vocal tract. An increase in tempo from normal to fast. results in consonant and vowel articulations following each other in quicker succession. Consonant gestures needing more execution time are therefore maintained, or even strengthened at the expense of articulatory gestures associated with vowel production. Consequently vowels tend to reduce in duration and quality, particularly if unstressed.

Although this may be the case for the languages cited above, it is not absolutely certain to what extent this could be classified as a universal of articulatory behaviour. Gay also stressed that the coordination of articulatory movements may be adjusted in some way to preserve the information-bearing elements across changes in tempo. For example, stressed vowel gestures may be maintained at the expense of unstressed vowel gestures. It is also possible that vowel gestures are only weakened if the phonology of the language allows it to happen. Certain languages tend to preserve vowel quality in unaccented or usrtressed position to a greater extent than others. As a loss of vowel quality is usually associated with durational shortening (Lindblom 1963), one has to consider to what extent phonology and mechanical inertia of the articulators interact under conditions of rapid articulation rate. Thus it seems likely that we need to put forward

language-specific rules of temporal organization that can account for physiological limitations, whilst maintaining "phonologically invariant temporal relations" (Port1974 p,272).

It is generally accepted that vowels in spoken French tend to preserve their distinctive quality to a much greater extent than vowels in languages such as English, under varying conditions of stress or accent. Delattre's (1968) comparative study of vowel reduction in French, German, English, and Spanish, describes in detail the tendency of English vowels to be modified acoustically and articulatorily, approaching more "centralised" qualities in unstressed position. French vowels on the other hand are modified vary slightly in comparison. This is often cited as a reason why spoken French gives the auditory impression of syllable timed whereas English sounds stressed timed. In spoken English stressed vowels may often be the only ones to sound fully distinct.

It is possible that given conditions of increased tempo, as well as presence or absence of accent, vowels in a French utterance may shorten until a certain limit is reached beyond which further shortening would result in a loss of informative acoustic cues to vowel identity, unless there was specific articulatory modification to preserve quality. This seems unlikely in view of the tendency of the speech production system to prefer hypo-articulation as opposed to hyper-articulation(Lindblom 1973). One might, nevertheless expect a smaller degree of vowel compression in French spoken at rapid tempo, than in other languages, like Dutch, Swedish or English, whose unstressed vowels reduce to short schwa-like segments. Furthermore, there may be a more equal degree of consonant and vowel duration compression in French than in these languages. Smith(1976)found this to be the case in her study. However the limitations of her corpora(data from one speaker) make it necessary to see whether this finding can be replicated in a much larger corpus of French.

Peterson and Lehiste(1960) and Port(1980) also investigated the interaction between tempo, and other duration influencing factors such as presence or absence of stress. Stressed syllables in many languages, including English, Dutch, and Swedish are generally much longer than unstressed syllables. Similarly accented syllables in French are also reportedly longer than unaccented syllables (Delattre 1968, Wenk and Wioland 1982). Peterson and Lehiste(1960) and Port(1960) found that stressed syllables in their American English data were less affected by rapid tempo increase than unstressed syllables. This is not surprising if it is interpreted in a stress-timing framework. If the basic "rhythm" of English is mainly governed by the alternation of stressed and unstressed syllables, it seems likely that unstressed syllables would be compressed to a greater extent than stressed syllables in order to maintain the rhythm in the light of increasing "articulatory demands" particularly if stressed syllables are somehow "more important" than unstressed syllables either from the articulatory, or perceptual point of view.

The interaction of tempo and accent in French has not to our knowledge been examined in any detail, so we do not know whether French behaves in similar ways to English. Proponents of the syllable- timing label of French would claim that there should be no difference at all given that this typology infers that accented syllables in French have no important role at the level of temporal organization. The basic unit of timing is said to be the syllable, whether accented or not. Consequently any re-organization of the articulatory plan due to increased articualtion rate should be evenly distributed across all syllables, and by extension their constituent segments. Thus the effects of tempo change on accented and unaccented vowel duration should be similar.

A handful of studies have examined the micro-effects of slowing down tempo in as much detail as speeding up tempo. There is a certain amount of consensus in the timing literature that slowing down and speeding up tempo are not converse processes. Port(1981) found that when subjects slowed down their neutral tempo, all consonant and vowel segments increased in duration by a constant ratio. Pickett(1980) reported similar findings. Smith (1976) reported a similar results for her French data; remembering also that she found no differences in the way consonant and vowel segment durations change as a function of tempo increase from normal to fast. Once again, the limitations of her experimental corpus make it necessary to test whether this finding can be replicated in a much larger data sample.

Experimental procedure

The following experiment was designed to investigate the above micro effects of tempo change in an extended corpus of spoken French. We predicted that there would be modifications in articulation rate between tempi. Furthermore, We wished in part, to test Smith's(1976) claim that there is no major difference in the degrees of consonant and vowel duration change between tempi (i.e. consonants and vowels compress or expand to a similar extent), due in part to the so-called syllable timing nature of French. On the basis of potential "braking" effects due to the tendency in French to maintain distinctive vowel quality in unaccented position, we hypothesized that the degree of consonant and vowel compression would not be as great between normal and fast tempo as the levels of consonant and vowel compression between normal and slow. We also hypothesized that if a strict syllable-timing description is relevant for French there should be no difference between the degree of shortening of accented and unaccented vowels from slow through to fast tempi.

Subjects and Materials

A transcript of a French radio interview was found that was not too long(112 words), that consisted of colloquial syntax and morphology, conceivably making it easy to read at either slow or fast tempo. Six native speakers of non-meridional French acted as informants in this experiment. All had been educated to university level. They were asked to read the text three times, once at their normal tempo, then at a slower rate, still maintaining natural delivery, and finally at a speeded up version of their normal tempo, likewise maintaining naturalness.

Analysis procedure

The informants' performance of this task was recorded under laboratory conditions on a FM Recorder (Racal Store 4D) at a speed of 30 ips. Oscillograms were made from these recordings using a Siemens-Elema Mingocard 4 at a paper speed of 25mm/s. The recordings had been slowed down by a factor of 4. This technique enables a more accurate segmentation of the audio trace into consonant and vowel segments.

Potentially informative changes in amplitude associated with the presence of nasal and lateral articulations, or intervals of aperiodic noise associated with frication are more readily identifiable. Certain segments such as semi-vowels were not easily identified on the waveform. These were analysed spectrographically, using a Kay Sona-graph. Duration measurements were obtained for intervals corresponding to consonant and vowel segments, following the procedures described by Peterson and Lehiste(1960) and Delattre(1965). Any intervals on the acoustic trace of 200ms and over, showing no signs of periodic or aperiodic noise above the level asociated with background noise were labelled pauses, following the procedure of Duez(1983). Syllables were identified according to phonolgical criteria, and their durations calculated by adding the durations of the constituant segments. In the case of elided segments or syllables, the segment or syllable were given the durational value 0.

The rate of articulation was measured according to the procedures oulined in Grosjean and Deschamps(1975), Butcher(1981),and Dauer(1983, and represented the number of syllables unttered per second of total articulation time. Durational data obtained from consonant and vowel measurements were tabulated according to tempo and speaker, and the means, standard deviations computed of segment duration (consonant and vowels combined), vowel and consonant duration independently, irrespective of accentuation and finally accented vowels and non-accented vowels. Withregards to the latter only vowel segments that occurred in all three readings for each speaker were included in the analysis. Analyses of variance were performed on segment duration in general, then consonant and vowel durations under the conditions of tempo change, and presence or absence of accent.

Results and Discussion

Tables 1-2 illustrate the results of these computations for each speaker. There seemed to be a consistant trend across all but one speaker to increase articulation rate from slow through to fast tempo. It was not evident from our results that articulation rate change was consistantly greater from normal to fast tempo than from normal to slow, as found by Butcher(1981) for German . However there was a significant degree of inter-speaker variability in determining each speakers fast, normal and slow tempi, as one would expect. Speaker 5 did not vary her articulation rate at all between normal and fast tempo, but reduced it by 35% between normal and slow.

We assumed that the above changes in articulation rate due to speeding up or slowing down tempo reflected a certain amount of segment compression and expansion. An examination of the means, standard deviations and results of the analyses of variance showed that segment durations were significantly affected by tempo for all speakers. There was a significant effect of Tempo across segments and subjects, although there was a certain amount of inter-speaker variability ($F = 111.01$, $P<.0001$). There was also a significant tempo/segment interacation ($F = 10.77$, $P<.0001$). This indicated that consonant and vowel segment durations were not interacting with tempo in the same way. An examination of the means and standard deviations of consonant and vowel duration highlight this difference. Vowel segment durations were compressed more than consonant segment dduration due to tempo increase. Thus our results do not replicate those of Smith(1976), and our French data seem to reflect similar patterns in this respect to the English data in Peterson and Lehiste(1960) and the Dutch data in Nooteboom and Slis (1972).

Speakers and	Tempo	TST(ms)	TAT(ms)	AT/TST	SR	AR	MDP	NP
1	*Fast*	38159	31528	83%	4.38	5.28	469(248)	12
	Normal	44611	33834	76%	3.76	5	790(542)	12
	Slow	55803	39020	70%	3.2	4.6	646(422)	21
2	*Fast*	26329	24673	94%	6	6.4	440(226)	4
	Normal	36082	29896	83%	4.65	5.6	478(368)	8
	Slow	47883	38818	77%	3.51	4.3	480(320)	16
3	*Fast*	30670	26900	88%	5.7	6.5	342(129)	8
	Normal	33471	29008	87%	5.2	6	372(195)	11
	Slow	40129	33573	84%	4.3	5.2	312(182)	12
4	*Fast*	30519	25959	85%	5.5	6.5	455(142)	10
	Normal	39684	32040	81%	4.3	5.3	507(241)	15
	Slow	57036	43033	75%	3	4	610(310)	23
5	*Fast*	36785	30270	82%	4.5	5.4	650(340)	10
	Normal	44138	30120	68%	3.7	5.4	825(554)	17
	Slow	87667	42000	62%	2.4	3.9	723(478)	25
6	*Fast*	25057	23387	94%	6.7	7.2	417(161)	4
	Normal	31781	28706	91%	5.3	5.9	514(173)	6
	Slow	38590	32994	84%	4.4	5.1	573(239)	11

Table 1. TST: total speaking time (ms); TAT: total articulation time (ms); AT/TST: the percentage articulation time to total speaking time; SR: speaking rate, number of syllables per second; AR: articulation rate (syllables per second); MDP: mean pause duration in ms (standard deviations in brackets); NP: number of pauses

Speaker	Tempo	Consonant (ms)	durations	Vowel (ms)	durations	Combined (ms)	segment
		mean	**standard deviation**	**mean**	**standard deviation**	**mean**	**standard deviation**
1	*slow*	92.18	(42.85)	102.47	(71.21)	96.92	(55.41)
	normal	81	(35.68)	91.16	(59.85)	85.68	(48.17)
	fast	72.7	(44.36)	75.72	(58.06)	74.09	(46.43)
2	*slow*	82.16	(38.99)	90.29	(49.33)	85.91	(34.84)
	normal	74.03	(36.04)	79.8	(52.94)	76.68	(36.00)
	fast	69.49	(34.56)	70.12	(38.71)	69.78	(33.75)
3	*slow*	84.08	(43.08)	93.05	(68.27)	88.22	(48.08)
	normal	78.43	(35.65)	76.61	(48.46)	77.59	(41.38)
	fast	69.53	(39.29)	62.80	(47.18)	66.43	(35.68)
4	*slow*	82.08	(33.07)	103.76	(84.54)	92.07	(55.08)
	normal	76.91	(49.45)	81.78	(60.09)	79.16	(46.61)
	fast	64.62	(25.94)	72.39	(61.7)	68.21	(35.66)
5	*slow*	90.73	(42.56)	108.18	(67.09)	98.77	(52.78)
	normal	81.9	(44.31)	76.84	(42.43)	79.62	(40.02)
	fast	83.14	(72.61)	74.18	(40.57)	79.01	(38.14)
6	*slow*	87.51	(53.64)	88.22	(62.14)	87.84	(50.29)
	normal	72.47	(36.85)	69.82	(41.14)	71.24	(38.47)
	fast	68.61	(32.67)	59.84	(30.74)	64.57	(33.26)

Table 2. Means and standard deviation values of consonant and vowel segment durations across tempi, and speaker.

The above result could be explained by the comparatively low values of articulation rate computed for these data. On the other hand Port's(1981) explanation of similar results in his data could account for our findings. That is, the similar degrees of compression from slow to normal and normal to fast could be explained by an interaction between physiological, mechanical needs of increasingly rapid articulation being offset by the need to maintain more or less invariant temporal relations determined by the phonology, or indeed rhythm, of the language in question.

Our results do not rule out the possibility that the need to maintain distinctive vowel quality in the fast readings could have been responsible for the paradigmatic relation between consonant and vowel duration compression between slow and normal, and normal and fast tempo. Although the proportion of vowel compression to consonant compression was greater from slow through to fast tempo, the fact that this proportion did not vary significantly between tempi seems at odds with results reported for English, for example. Pickett(1982) found that consonants and vowels in his data were expanded by 33% from normal to slow tempo, but the ratio of vowel compression to consonant compression was much greater from normal to fast tempo. however further tests need to be carried out on our data before we can draw any further conclusions.

Comparing the effects of accent and tempo, there was once again a significant degree of inter-speaker varibility withregards to segment duration ($F=13$; $P<0.0000$). The effects of accent were also strongly significant ($F=243$;$P<0.0000$). The interaction between accent, tempo and subject was extremely pronounced ($F=3.59$;$P<0.0001$). Thus vowel duration in all of the readings was substantially affected by accent and tempo.

A further analysis of variance and covariance was performed on the vowel duration data to test whether there were differences in the way accented and accented vowels were affected by tempo change. This proved to be the case. Tempo affected stressed and unstressed vowels in different ways (F=12.33; P>0006). However unlike the results of Peterson and Lehiste(1963) for American English, and Nooteboom and Slis(1972) for Dutch, tempo change affected accented vowel duration more substantially than unaccented vowel duration. Moreover, there were significant differences in the degree of duration change between slow and normal tempo and normal and fast tempo. The mean differences for accented vowel duration between slow and normal tempo were of the degree of 37%, whereas between fast and normal tempo they measured 16%. However the values for unaccented vowels, 11.5% and 10.95% between slow and normal tempo and fast and normal tempo, were not only a good deal smaller than for accented vowels, but there was no significant difference between tempi.

Speaker	Tempo	Accented Vowel Duration (ms)		Unaccented Vowel Duration (ms)	
		mean	standard deviation	mean	standard deviation
1	*slow*	172.23	73.95	78.21	26.45
	normal	142.18	70.04	66.47	20.80
	fast	129.3	89.43	56.32	17.11
2	*slow*	146.97	45.04	70.61	19.4
	normal	128.33	49.38	62.09	19.67
	fast	120.25	45.3	54.68	18.9
3	*slow*	170.41	82.58	67.87	28.10
	normal	131.66	50.04	55.05	21.67
	fast	103.46	44.02	45.42	14.61
4	*slow*	182.95	84.28	69.31	19.08
	normal	142.18	59.2	58.49	17.87
	fast	120.38	51.36	51.74	15.67
5	*slow*	175.8	69.02	81.1	22.58
	normal	134.64	48.76	66.56	19.54
	fast	120.66	40.33	59.16	19.48
6	*slow*	164.36	76.09	67.36	18.85
	normal	112.18	47.49	56.63	18.04
	fast	97.25	35.18	45.89	15.85

Table 3: Means and standard deviation values of accented and unaccented vowel duration across tempi and speakers.

There are a number of possible reasons why this pattern emerges in our results. The mean durations for unaccented vowels are a great deal shorter than the mean duration of accented vowels, in general. Consequently the smaller tempo effects on unaccented vowel duration could be due to a compressibility limit operating to prevent over-shortening that may lead to loss of distinctive vowel quality. Furthermore, Smith (1976) found that longer segments in her data(whether consonants or vowels) tended to be shortened a great deal more than intrinsically short segments. However, we have not at this stage taken into account the interaction between phonological or intrinsic consonant or vowel length, tempo and accent. A further analysis of the data bearing this factor in mind is therefore necessary.

More significantly, the large difference between accented vowel duration in the slow and normal readings is more than likely concomitant with the reduction in the number of pauses between the three readings(see table 1). Given that accented syllables in French are largely group final, they are often pre-pausal, and subsequently are longer than non-pre-pausal syllables(Smith 1976). Therefore a reduction in the number of pauses between readings necessarily reduces the number of pre-pausal syllables. we are in the aprocess of investigating this possibility.

Smith(1976) claims that the only truly "long" segments in French are those belonging to pre-pausal syllables. But as accent and the possible associated durational properties were not taken into account in her study, it remains unclear if or how non-pre-pausal accented syllables interact differently with tempo than pre-pausal accented syllables. On the basis of our finding that the margin of difference between the amounts of shortening of accented and unaccented vowels was substantially reduced between normal and fast tempo as opposed to slow and normal tempo, we could hypothesise that if pre-pausal vowels were excluded from the analysis the margin of difference would disappear.

Finally, we did not take into account phenomena such as vowel and syllable elision reported to occur as a result of increasing tempo from normal to fast (Smith 1976). Comparative instrumental studies of the micro-effects of tempo increase in French and English are non-existant. Consequently we do not know for sure whether French relies more on syllable and segment elision than segment compression to increase tempo. It is also uncertain whether the number of accented syllables is reduced to counteract overshortening of individual segments. If both of the above possibilities are in fact the case, this could explain why our results differ from those reported for other languages. Although the number of accented syllables in our data did vary, there was some indication that there was some kind of resistence to eliminate too many. A strict syllable-timing hypothesis would predict a reduction in the number of accented syllables. It seems that stressed syllables in English almost never disappear under conditions of rapid articulation rate associated with tempo increase. Once again we need to examine the data in more detail before we can make any further assumptions about the differences between the two languages, and whether we can involve the syllable-timing and stress-timing dichotomy.

Conclusion

It certainly appears to be the case that there are some language- independent ways in which the acoustic durations of a stretch of connected speech reflect tempo change, and in particular tempo increase. The patterns of segment compression in French due to tempo increase are more complex than reported previously. Like English, Dutch or Swedish, vowel duration is more strongly affected by changes in speech rate associated with tempo change, than consonant duration. However it seems that at this stage there may be strong differences between the degree of duration change experienced by different languages.

The situation is much less clearcut when it concerns the interaction between vowel duration, positional and accentual variables, and tempo change. So little experimental data exists to show how tempo interacts with the above duration influencing variables in any language. It seems reasonable at this stage to acknowledge that French accented syllables do not behave in the same way as English or Dutch stressed syllables. Whether this is due to the so-called syllable/ stress timing dichotomy is not absolutely clear. There are definite differences in the way accented and unaccented vowels respond to tempo change in French. A simplistic typology label such as syllable-timing cannot account for these differences.

However, it seems likely that any differences between French and the above-mentionned languages is due to not so much the behaviour of accented vowels, but unaccented vowels. That is to say the degree of difference is not so much due to accented vowels not resisting durational compression unlike unaccented or unstressed vowels in Dutch or English. The reverse seems to operate in our corpora of French. The smaller degrees of temporal compression of unaccented vowels reflect a difference between French and these other languages. The reasons for this compression resistance include the tendency to maintain intrinsic vowel quality in unaccented syllables in spoken French. Further analyses of the data along the guidelines suggested in the above sections are currently being undertaken. Nevertheless we may interpret our present results as evidence of the trade-off proposed by Port (1981) between the mechano-inertial demands of rapid articulation as part of the strategies to increase speaking tempo and phonological conditioning.

References

Butcher, A.R. (1981): Aspects of the speech pause: phonetic correlates and communicative functions. Phd Thesis, Universitaet Keil. Also as: Arbeitsberichte des Instituts fuer Phonetik der Universitaet Kiel 15.

Delattre, P. (1968): An acoustic and articulatory study of vowel reduction in four languages. IRAL VII/4, 295-325.

Gay, T. (1978): Effect of speaking rate on vowel formant movements. JAcoustSocAm 63, 223-230.

Lindblom, B. (1963): Spectrographic study of Vowel reduction. JAcoustSocAm 5, 1773-1781.

Lindblom, B. (1983): Economy of speech gestures. In Macneilage,P.F.(ed.) The Production of Speech. New York, Heidelberg, Berlin: Springer-Verlag.

Miller, J.L., Grosjean, F. and Lomanto, C. (1984): Articulation rate and its variability in spontaneous speech: a reanalysis and some inplications. Phonetica 41, 215-225.

Nooteboom, S.G. (1972): The interaction of some intra-syllabic and extra- syllabic factors acting on syllable nucleus durations. IPO AnnProgRep 7, 30-39

Peterson, G.E. and Lehiste, I. (1960): Duration of syllabic nuclei in English. JAcoustSocAm 32, 693-703.

Pickett, J.M. (1982): The Sounds of Speech Communicatioon: A Primer of Acoustic Phonetics and Speecn Perception. Baltimore: University Park Press.

Port, R.F. (1981): Linguistic timing factors in Combination. JAcoustSocAm 69, 262-274.

Smith, A. (1976): The timing of French, with reflections on syllable timing. Work in Prog Dept Ling Edinburgh Univ 9, 97-108.