A MICROPROSODIC APPROACH TO ANALYZING SPEECH RATE EFFECTS IN SOCIOPHONETIC VARIATION

Jeff Tennant
University of Western Ontario, London, Ontario, Canada

ABSTRACT

This paper outlines a method, based on a normalization procedure using microprosodic criteria, for investigating speech rate effects on deletion variables in French. Weighting factors are established based on the intrinsic length of segments (vowels, consonants, and glides) as well as the conditioning effect of phonological environment (stress, syllabic position), as observed in a corpus of spontaneous speech. These weighting factors are used to adjust duration measurements taken in the immediate context of each token of the dependent variable, in order to obtain a reading for speech rate (in segments per second) which allows comparison across samples differing in length and phonetic content. The method is applied to L-deletion in the pronoun *elle* in the speech of a Canadian French speaker. Comparison of mean normalized speech rates for tokens with L maintained and those with L deleted shows a significant effect: L is deleted more frequently at higher speech rates.

1. INTRODUCTION

The purpose of this paper is to present a method for investigating, in spontaneous speech, what effects speaking rate might have on variables involving the deletion of segments. The primary concern will be with describing an approach which is proposed as a means of solving the methodological problems which stand in the way of adequately quantifying these effects in variation studies. I will illustrate this approach by analyzing a variable sometimes assumed to be conditioned by speech rate: deletion of L in the subject pronoun *elle*, in the idiolect of one speaker of Canadian French.

Retention or deletion of a phonetic segment characterizes many variables which have attracted the attention of sociolinguists, going back at least as far as post-vocalic R-deletion in Labov's [10] New York City department store survey. A considerable body of research carried out since the 1960s demonstrates what social and linguistic factors condition variables such as presence or absence of final T and D in English, or presence or absence of the consonant L in Canadian French pronouns and articles. Guy [6], Sankoff & Cedergren [17] and Poplack & Walker [16] are only a few examples. Authors of studies on some of these variables have occasionally raised the possibility that how quickly a person speaks might also be a relevant factor in these deletion variables. For example, Santerre, Noiseux and Ostiguy [18] hypothesize that differences in speaking rate would explain the differences observed between older and younger speakers in frequency of L-deletion in Montréal French. Similarly, Guy [6] comments in passing that speaking rate likely plays a role in final stop deletion in American English, but he does not include it in his quantitative study due to methodological limitations.

The assumption that speaking rate and deletion variables would be correlated in this way may appeal to common sense intuition, and is in all likelihood well grounded in folk linguistic beliefs. Jack Chambers (personal communication) views it as a "pedestrian claim" that probably needs to be seriously qualified. But we do not have enough information yet to be certain about how well founded this assumption is in empirical fact, or about which variables are affected by speaking rate and which ones are not. Shockey's [20] remarks are still valid today: "In a sense, we currently have a theory about how to construct a set of algorithms which can be used to generate the phonological surface forms found in fast or conversational speech, but we cannot actually construct them due to lack of information." (p. 217)

The hypothesis that a faster rate of speech will lead to a higher frequency of surface reductions can be formulated at two distinct levels, which I will refer to here as the global level and the local level. On the global level, the hypothesis can be formulated as follows: a speaker who generally speaks quickly will have a higher rate of deletion (of L in the French subject pronoun *elle*, for example) than a speaker who habitually speaks slowly. On the local level, the hypothesis can be formulated somewhat differently. For a variable whose two variants are deletion and retention of a segment, in the speech of a given speaker who uses both variants, occurrences of deletion will be found in contexts where speaking rate is more rapid. Similarly, tokens in which L is pronounced will be found in contexts where speaking rate is slower.

This paper will concentrate on the second of these hypotheses, which is more difficult to test than the first one. I will present and discuss the results of an analysis using a procedure for normalizing speaking rate measurements obtained in the immediate context of each occurrence of the dependent variable. The principle of this normalization method involves the application to segment duration measurements of what Santerre & Roberge [19] term "microprosodic weighting factors". The dependent variable analyzed is the deletion of the consonant L in the French subject pronoun *elle*, a variable which has been studied numerous times since Sankoff & Cedergren's [17] study, but for which the effects of speaking rate have not yet been conclusively demonstrated.

2. PREVIOUS STUDIES

2.1. Experimental studies

Speaking rate effects on deletion variables have been examined in a small number of experimental studies in which speakers are recorded while reading utterances aloud. For example, in Shockey's [20] study, participants are asked to read a series of sentences using two different rates of speech. The results show that T/D deletion increases when subjects speak faster. Trépanier and Archambault [24], using a similar method, observe the same tendency for high-vowel deletion in Quebec French.

2.2. Studies using spontaneous speech

Studies of rate effects on deletion carried out using spontaneous speech samples are even fewer in number. Léon [12] observes that speaking rate has a slight influence on schwa-deletion in French, but he does not elaborate on the methodology used to measure rate. In two studies I have carried out based on analyses performed on a corpus of adolescent speakers Ontario French [22, 23], some confirmation for the hypothesis of rate effects on reduction was found. More specifically, it was found that "global" measures of speech rate and other temporal variables [5] appeared to correlate with frequency of L-deletion in definite articles for four speakers [22]. In another study using "local" speech rate measurements, higher speaking rates appeared to favor a greater frequency of L-deletion in the subject pronoun *elle* [23]. In that study, I began to address the methodological problems involved in using localized measurements of speech rate, but was far from satisfied with the solutions I had proposed. The goal of this paper is to continue the study of these problems, and to present what I believe is a solution which can be refined to achieve greater accuracy in the analysis of speech rate effects.

3. CORPUS AND METHODOLOGY

3.1. Speaker

The corpus used for this study is a 45-minute interview with a single speaker, Diane, a middle-class bilingual French-Canadian woman from Welland, Ontario. At the time the corpus was gathered by Raymond Mougeon's research team in 1974, Diane was 21 years old.

3.2. Signal analysis

All rhythmic groups containing occurrences of the dependent variable *elle* were digitized from the recording using SoundEdit 16 on a Power Macintosh. Rhythmic group boundaries were determined based on the researcher's perception of sequences of unstressed syllables with a stressed syllable at the end.

The beginning and end points of each phonetic segment were marked in the digitized signals by means of the label function in the Signalyze speech analysis program. Information from the oscillographic representation as well as spectrograms and amplitude envelopes was used to perform the segmentation. These label files were imported into Excel spreadsheets, which allowed efficient and verifiable calculation of segment duration, as well as application of the normalization algorithm.

This part of the corpus contains 68 tokens of *elle*, where the pronoun is unambiguously a subject clitic. False starts and repetitions involving this pronoun were excluded, as were occurrences of stressed *elle*. The 68 rhythmic groups containing these occurrences of *elle* contain a total of 541 phonetic segments.

3.3. Normalization

The normalization method used here is based on a microprosodic weighting system described by Santerre & Roberge [19]. The purpose of the procedure is to obtain accurate measurements of speech rate in the immediate context of the dependent variable under study by adjusting duration readings to take into account factors of intrinsic and conditioned duration in segments.

As Santerre & Roberge [19] point out, acoustic measurements of duration need to be normalized to eliminate the effects of mechanical or physiological production in order to be taken as true representations of temporal variations controlled by the speaker. One might also add that, for our purposes here, there is also a need to compensate for the effects of some factors, such as stress, which are linguistically manipulated by the speaker. The major problem being addressed here concerns comparing speaking rate measurements based on samples of usually two to seven syllables which do not have the same phonetic segments or the same syllable structures. Without some kind of normalization procedure, such comparisons are not likely to be accurate.

A number of factors can influence the duration of a phonetic segment in French [1, 2, 3, 4, 9, 11, 13, 14, 15] and their effects can be divided into intrinsic and conditioned (or co-intrinsic) duration differences. Intrinsic duration differences in vowels include place and manner of articulation (aperture, backness, nasality) and the presence of phonologically long vowels, particularly in North American dialects. For consonants, place of articulation and voicing account for most intrinsic duration differences. Conditioning factors on vowel duration include stress and the nature of preceding onset and especially the following coda. Consonant durations are partially conditioned by stress, syllabic position and the place and manner features of adjacent consonants in clusters.

In order to obtain accurate local readings of speech rate, we need a procedure which will compensate for these intrinsic and conditioned divergences among segment durations. The approach proposed here is conceived in terms of adjustments in relation to an average segment length.

Santerre & Roberge [19] present a series of weighting factors which were calculated using a corpus of Quebec French speech elicited by means of a reading task. The applicability of these factors is somewhat limited, since they represent segment durations in CVC syllables only and they exclude sonorant consonants and glides. In addition, it would be advantageous at this stage, in order to assure accuracy, to establish weighting factors based on a sample of speech from the speaker whose use of the linguistic variable is being studied.

To this end, a sample consisting of 2 1/2 minutes of continuous speech from Diane's interview was used to establish weighting factors. In this sample, the duration measurements for phonetic segments were made as they were for the tokens of *elle*. Linked spreadsheets were used to calculate average durations of each segment, taking into account the intrinsic and conditioned factors which Santerre & Roberge showed to have the greatest effects on segment length. These average values were then converted to proportions in relation to a single reference point, the unstressed vowel /e/ in open syllables. These proportions, shown in Tables 1 and 2, represent factors by which each segment duration reading is to be multiplied in order to obtain a normalized figure. This normalized figure therefore corrects for divergences from an idealized mean segment length, which could be envisaged in terms as a slot on the CV tier in an autosegmental phonology model. These normalized duration measurements are used to calculate the speaking rate corresponding to each occurrence of the pronoun *elle* in segments per second.

	With Primary Stress			No Primary Stress		
	open syll	closed syll	closed + v z ʒ vr	open syll	closed syll	closed + v z ʒ vr
i	0.76	0.82	0.64	1.22	1.06	1.19
y	0.59	X	X	1.42	1.00	1.29
u	X	X	0.50	1.06	1.76	1.74
e	0.57	X	X	1.00	1.28	X
ø	0.69	X	X	1.20	X	X
o	0.64	0.59	X	0.76	1.42	0.59
ɛ	X	X	X	1.05	X	X
ɛ	0.66	0.78	0.57	1.13	1.70	1.35
œ	X	0.78	X	1.17	1.57	0.71
ɔ	X	X	X	1.68	1.48	1.30
a	0.88	0.92	0.43	1.20	1.22	1.53
ɑ	1.07	0.28	0.50	1.03	X	X
ɛ:	0.37	X	X	X	0.73	X
ɛ̃	0.71	X	X	0.81	X	X
œ̃	0.76	0.79	X	0.94	X	X
ɔ̃	0.82	X	X	1.54	X	X
ɑ̃	0.65	0.64	X	0.82	0.89	X

Table 1. Weighting factors for adjustment of vowel durations. X = No occurrences of segment in this context in corpus.

	With Primary Stress			No Primary Stress		
	simple onset	complex onset	coda	simple onset	complex onset	coda
p	0.75	0.97	X	0.88	1.20	1.25
t	0.80	0.94	1.09	0.90	0.96	1.27
k	0.67	0.81	0.97	1.06	1.19	1.90
b	X	1.80	X	1.20	1.00	X
d	0.91	0.96	2.31	1.34	X	1.53
g	X	X	X	X	2.00	X
f	0.62	X	X	0.74	1.02	X
s	0.71	0.68	X	0.79	1.01	0.94
ʃ	X	X	1.30	X	X	X
v	1.39	1.03	X	1.40	1.47	X
z	0.96	0.93	0.86	1.49	1.64	X
ʒ	1.00	X	0.62	1.33	1.03	1.14
m	0.88	1.47	0.93	1.24	0.97	1.59
n	1.28	1.80	0.76	2.12	X	0.92
ɲ	2.00	X	X	X	X	X
ŋ	X	X	X	X	X	X
l	1.43	1.22	1.40	1.61	1.56	1.81
ʀ	0.64	1.55	1.29	1.15	1.50	1.61
r	1.66	2.39	X	2.16	1.99	2.24
ɥ	X	X	X	1.35	1.57	X
w	0.88	1.98	X	1.44	1.83	X
j	1.68	2.05	0.95	2.18	1.83	1.06

Table 2. Weighting factors for adjustment of consonant durations. X = No occurrences of segment in this context in corpus.

4. RESULTS AND DISCUSSION

This method was applied to the 68 occurrences of the subject pronoun elle in Diane's interview. Table 3 gives overall deletion rates as well as deletion rates according to the following phonological context.

Following context	L pronounced	L deleted	N	% deletion
/_C	22	19	41	46.3%
/_V	14	13	27	48.1%
Total	36	32	68	47.1%

Table 3. L-deletion in the subject pronoun *elle* for one French Canadian speaker.

Interestingly, Diane does not follow the trend observed in other studies [16, 17] where L tended to be deleted more frequently before consonants than before vowels.

NO NORMALIZATION	Average (seg/sec)	stand. dev.
L pronounced	12.57	3.45
L deleted	14.07	3.10
t-test	p =	0.063

Table 4. Speaking rates for tokens of L with L pronounced and L deleted. Rate measurements not normalized.

Table 4 shows the averages and standard deviations for speaking rates corresponding to retention and deletion occurrences of the variable, without normalization. Although the direction of the difference between speaking rates confirms the hypothesis that deletion is more frequent at higher speaking rates, the t-test shows that the difference between these means in not significant even at the $p < .05$ level.

NORMALIZED DATA	Average (seg/sec)	stand. dev.
L pronounced	11.70	3.08
L deleted	14.08	3.18
t-test	p =	0.003

Table 5. Speaking rates for tokens of *elle* with L pronounced and L deleted. Rate measures normalized using weighting factors.

The data in Table 5 represent speaking rates for the same occurrences of *elle*, but this time with duration measurements adjusted using the microprosodic weighting factors shown in Tables 1 and 2. These results indicate that the mean speaking rate for occurrences where the L is deleted is greater than the mean rate for occurrences where L is retained, by a difference of 2.38 segments per second. Although the rather high standard deviations suggest that this result must be interpreted with caution, the t-test does indicate that this tendency is significant at the $p < .01$ level. There is therefore clear evidence that, for this speaker, faster speaking rates correspond to higher rates of L-deletion in *elle*.

What are the linguistic implications of such a result? Solé [21] suggests that the presence or absence of speech rate effects can contribute to distinguishing between phonologized low-level phonetic processes (which should vary with respect to speech rate) and lexicalized phonetic/phonological processes (which should be rate invariant). However, other studies of L-deletion, summarized by Poplack & Walker [16], tend to situate this variable at a morphophonemic level, i.e. at a level which makes reference to phonological information which is represented lexically in individual morphemes. A model such as Guy's variable lexical phonology [7, 8] could perhaps offer a useful explanatory framework for this apparently rate-dependent deletion process. A full linguistic analysis is, however, beyond the scope of this paper.

The method presented here still needs refinement. By establishing microprosodic weighting factors based on a larger sample of speech, it will be possible to take into account other parameters not studied here, such as the effect of the onset consonant on the following vowel. Treatment of stress also needs to be improved. Different degrees of stress (primary and secondary) should be taken into account, and the Canadian French phenomenon of vowel lengthening in the penultimate syllable needs to be better accounted for. Finally, a refined system of weighting factors based on a larger data sample could also take into account the effects of specific phonemes on the duration of adjacent segments, as well as less frequent phenomena such as consonant durations in clusters, both in onset and coda positions.

The system outlined here is of course specific to French. Much different configurations of factors would need to be elaborated for other languages, taking into account the prosodic particularities of each language [3]. Nonetheless, there is little doubt that the same underlying principles apply.

5. CONCLUSION

This study presented a method for normalizing speech rate readings for the study of their effects on deletion variables in spontaneous speech. Although this method still needs refinement, it does compensate for many of the factors responsible for the segment duration differences which would otherwise render localized rate measurements inaccurate. The application of this method to L-deletion in Canadian French subject pronouns lent support to the hypothesis that deletion is affected by rate. As Solé [21] argues, speech rate data can provide valuable evidence for determining the linguistic nature of processes which could be operating at the phonetic, phonological or lexical level, or at more than one of these levels.

Guy [6] points out that, at the time of his T-/D-deletion study, a simple and reliable system was not available for coding speech rate. The method presented here is not a simple one, and will probably have to be made more complex in order to be fully reliable. But that is the nature of the phenomenon under study here, and progress in empirical knowledge in the area will require a considerable amount of effort to be devoted to corpus analysis.

ACKNOWLEDGEMENTS

This research is funded by the Social Sciences and Humanities Research Council of Canada (Standard Research Grant #410-96-0265). The author wishes to thank the following people for their comments on previous versions of this work: Parth Bhatt, Jack Chambers, Wladyslaw Cichocki, David Heap, Dennis Preston.

REFERENCES

[1] Boudreault, M. 1968. *Rythme et mélodie de la phrase parlée en France et au Québec*. Québec: Presses de l'Université Laval.

[2] Delattre, P. 1959. Rapports entre la durée vocalique, le timbre et la structure syllabique du français. *The French Review*, 22, 6, 547-552.

[3] Delattre, P. 1965. *Comparing the Phonetic Features of English, French, German and Spanish, an Interim Report*. Heidelberg: Julius Groos Verlag.

[4] Di Cristo, A. 1980. La durée intrinsèque des voyelles du français. *Travaux de l'Institut de Phonétique d'Aix*, 7, 311-235.

[5] Grosjean, F. 1980. Temporal variables within and between languages. In Dechert, H. and Raupach, M. (ed.), *Towards a Cross-Linguistic Assessment of Speech Production*. Frankfurt: P. Lang, 39-53.

[6] Guy, G. 1980. Variation in the group and in the individual: the case of final stop deletion. In Labov, W. (ed.), *Locating Language in Time and Space*. New York: Academic Press, 1-36.

[7] Guy, G. 1991a. Explanation in variable phonology: An exponential model of morphological constraints. *Language Variation and Change*, 3, 1-22.

[8] Guy, G. 1991b. Contextual conditioning in variable lexical phonology. *Language Variation and Change*, 3, 223-239.

[9] Jacques, B. 1974. Variations de durée des voyelles et des consonnes fricatives post-vocaliques finales de syllabe en position accentuée et inaccentuée. *Cahier de linguistique*, 4, 89-115.

[10] Labov, W. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

[11] Laver, J. 1994. *Principles of Phonetics*. Cambridge: Cambridge University Press.

[12] Léon, P. 1987. E caduc: facteurs distributionnels et prosodiques dans deux types de discours. *Proceedings of the XIIth International Congress of Phonetic Sciences*. Academy of Sciences of the Estonian S.S.R., vol. 3, 1-4.

[13] O'Shaughnessy, D. 1981. A study of French vowel and consonant durations. *Journal of Phonetics*, 9, 385-406.

[14] O'Shaughnessy, D. 1984. A multispeaker analysis of durations in read French paragraphs. *Journal of the Acoustical Society of America*, 76 (6), 1664-1672.

[15] Ouellet, M. 1992. *Systématique des durées segmentales dans les syllabes en français de France et du Québec*. Ph.D. Thesis. Université de Montréal.

[16] Poplack, S. and Walker, D. 1986. Going through (L) in Canadian French. In Sankoff, D. (ed.), *Diversity and Diachrony*. Amsterdam: John Benjamins, 173-198.

[17] Sankoff, G. and Cedergren, H. 1971. Les contraintes linguistiques et sociales de l'élision du L chez les Montréalais. In Boudreault, M. and Moehren, F. (ed.), *Proceedings of the XIII International Congress of Romance Linguistics and Philology*. Québec: Presses de l'Université Laval, 1101-1116.

[18] Santerre, L., Noiseux, D. and Ostiguy, L. 1977. La chute du /l/ dans les articles et les pronoms clitiques en français québécois. *The Fourth LACUS Forum 1977*, 530-538.

[19] Santerre, L. and Roberge, M. 1992. Facteurs de pondération psychoacoustique des durées en fonction de la nature des segments syllabiques et de l'accentuation en français du Québec. In Martin, P. (ed.), *Mélanges phonétiques et phonostylistiques offerts au professeur Pierre Léon*. Toronto: Éditions Mélodie, 439-461.

[20] Shockey, L. 1987. Rate and reduction: Some preliminary evidence. In Channon, R. and Shockey, L. (ed.), *In Honour of Ilse Lehiste*. Dordrecht-Providence: Foris Publications, 217-224.

[21] Solé, M.-J. 1995. New ways of analyzing sound change: Speech rate effects. *Belgian Journal of Linguistics*, 9, 21-44.

[22] Tennant, J. 1993. Le débit de la parole peut-il avoir un effet sur la variation morphonologique? In Ferguson, B., Gezundhajt, H. and Martin, Ph. (ed.), *Accent, intonation et modèles phonologiques*. Toronto: Éditions Mélodie, 35-54.

[23] Tennant, J. 1995. Dealing with Speaking Rate as a Factor Group in Variation Studies. Paper presented at NWAV(E) 24, University of Pennsylvania.

[24] Trépanier, C. and Archambault, D. 1992. Chute de segments et traces de durée. *19èmes J.E.P.* Bruxelles, 19 au 22 mai 1992, 59-64.