Phonetica *41:* 215–225 (1984)

0031–8388/84/0414–0215$2.75/9

Articulation Rate and Its Variability in Spontaneous Speech: A Reanalysis and Some Implications

Joanne L. Miller, François Grosjean, Concetta Lomanto

Northeastern University, Boston, Mass., USA

Abstract. It is by now well established that during normal conversation talkers often produce large variation in the rate at which they speak. However, existing research suggests that this modification is largely due to changes in the amount of pausing during conversation, and much less to actual changes in articulation rate, that is, the rate at which the speech itself is produced. In an attempt to examine this issue further, we used a modified measurement procedure to reanalyze the speech data from 30 talkers in an interview situation. In contrast to the earlier analyses, we found that there was indeed substantial variation in articulation rate for these speakers, even within a single utterance of a single talker. The implications of these findings for theories of segmental perception and for models of speech planning are discussed.

Introduction

A primary goal of research on speech perception is to specify the mapping between the acoustic signal of speech and its phonetic structure. Over the years, it has become increasingly clear that this mapping is complex. One aspect of this complexity is that many of the properties of the speech signal that convey segmental information typically vary as a function of a number of factors including, for example, speaker, phonetic context, stress pattern, intonation contour, and speaking rate. Indeed, a major goal of research has been to specify precisely how the acoustic form of speech varies with these factors and to explicate the manner in which the listener accommodates for this variation in the course of deriving the phonetic structure of an utterance [for reviews of this literature, see *Jusczyk,* 1984; *Liberman* et al., 1967; *Pisoni,* 1978; and for a different perspective on the problem, see *Stevens and Blumstein,* 1981].

One source of variation that has recently received particular attention is speaking rate. Numerous studies have shown that in a laboratory setting speakers are able to produce specified utterances (e.g. single words, simple phrases) at varying rates of speech [*Port,* 1981; *Tuller* et al., 1982], and acoustic analyses of utterances produced in such controlled experimental situations have shown that the alterations in rate modify the acoustic fine structure of individual syllables, affecting many of those properties that convey segmental information, for both consonants and vowels [*Port,* 1981; *Summerfield,* 1975]. Moreover, perceptual research has revealed that listeners are extremely sensitive to these variations and that, during the course of speech processing, they treat the segmentally relevant acoustic properties in a rate-dependent manner [*Summerfield,* 1981]. That is to say, during the course of identifying the segmental structure of speech, listeners take into account the rate at which the speech was produced [*Miller,* 1981; *Nooteboom* et al., 1978; for pertinent reviews].

The significance of this research for theories of segmental perception derives in part from the belief that changes in rate occur not only in a laboratory setting, but also in the course of normal conversation. The critical assumption is that speakers in a spontaneous setting change their rates considerably, with concomitant changes in the acoustic fine structure, such that the type of rate-dependent processing demonstrated in perceptual experiments plays an important, perhaps necessary role in natural language understanding. However, the existing spontaneous production literature would appear to call this assumption into question. Many studies have shown that overall speaking rate – typically measured in words or syllables per minute – does indeed vary considerably, depending on such factors as the individual speaker, the emotional state of the speaker, the type of speaking situation, the familiarity of the material being discussed, and the like [*Goldman-Eisler,* 1968; *Grosjean and Deschamps,* 1973, 1975]. Overall speaking rate is a complex variable, however, as it is composed both of the rate at which the speech itself is produced – articulation rate – and the number and duration of pauses in the utterance – pause rate. And it has been well established that the overall changes in speaking rate that occur both within and across speakers are largely due to changes in the pause rate, with the articulation rate remaining relatively stable across changes in conditions and within a single conversation [*Goldman-Eisler,* 1956; *Grosjean,* 1980]. From these findings one could conclude that the changes in overall rate that occur during conversational speech do not greatly alter the acoustic structure of the speech itself, calling into question the need for rate-dependent segmental processing.

But this conclusion is premature, in that two issues have failed to be addressed in the literature: the magnitude of variation issue and the measurement issue. The first issue concerns what constitutes significant variation. Although the changes in articulation rate that occur during conversation may indeed be considerably less than the changes in pause rate, they may still be quite substantial with respect to the durational characteristics of the acoustic properties specifying phonetic structure, which are often in the range of tens of milliseconds.

Consider the three investigations that have, to our knowledge, provided the most detailed analyses of speaking rate in conversational speech, beginning with that of *Goldman-Eisler* [summarized in *Goldman-Eisler,* 1968]. Over a number of years, *Goldman-Eisler* conducted extensive studies on the temporal structure of speech produced in a variety of situations, including psychiatric interviews, collegial discussions, and descriptions and summaries of cartoons. Regarding speaking rate, her major conclusion was that both across speakers and situations, articulation rate remained relatively stable compared to pause rate. But consider the magnitude of changes in articulation rate that were reported in a representative study [*Goldman-Eisler,* 1956]. The speech from 8 persons being interviewed in a psychiatric setting was measured; this included 5 patients and 3 staff members. For each interview, the speech was divided into utterances, with an utterance defined as the speech between successive questions of the interviewer. For each utterance, the number of syllables, speech time, and pause time were determined, and articulation rate was computed as the number of syllables per time speaking (discounting time pausing). Thus, one articulation rate was computed per utterance. The mean articulation rate for a given speaker, averaged across utterances of that speaker, ranged from 4.4 syllables/s for the slowest speaker to 5.9 syllables/s for the fastest speaker. In terms of the average duration of a syllable within an utterance, this constitutes a range between 227 and 169 ms. Especially important for our purposes is the standard deviation in rate across the utterances of a given speaker; this value ranged from 0.54 to 1.48 syllables/s, yielding a coefficient of variation ranging from 11.5 to 25.1%. Thus, at least for some speakers, we can conclude that there was reasonable variation in the articulation rate per utterance across the interview situation and this despite *Goldman-Eisler's* [1968, p. 25] conclusion that '... rate of articulation is a personality constant of remarkable invariance'.

A second major examination of speaking rate was conducted by *Malécot* et al. [1972], as part of a larger study of numerous acoustic-phonetic characteristics of conversational French. The corpus consisted of 50 half-hour conversations with members of the Parisian 'establishment'. The articulation rate of each utterance in the corpus was determined; in this study, an utterance was defined as a stretch of speech, two syllables or greater in length, between two pauses. The mean articulation rate across all utterances in the corpus was 5.73 syllables/s – the lowest articulation rate was 1.67 syllables/s and the highest was 9.67 syllables/s, with a standard deviation of 0.83 syllables/s, or a coefficient of variation of 14.5%. Recomputed in terms of average time per syllable within an utterance, the mean rate is 175 ms/syllable, with a range from 599 ms/syllable for the slowest utterance in the corpus to 103 ms/syllable for the fastest utterance in the corpus. Although these statistics do not allow us to determine the extent of variation in rate for a single speaker (and the data for individual speakers were not reported), they clearly suggest that average syllable durations in conversational speech are quite variable.

Finally, we turn to the study of *Grosjean and Deschamps* [1975], which forms the basis of the present investigation. These investigators examined conversational speech from radio interviews with 30 French and 30 English speakers. A main finding was that for both languages, changes in overall speaking rate were due primarily to changes in pausing, with considerably smaller changes in articulation rate. But the changes in articulation rate were not negligible. Across speakers and 'utterances' (defined below), the average articulation rate for French was 5.29 syllables/s with a standard deviation of 0.40, yielding a coefficient of variation of 7.5%, and for English it was 5.17 syllables/s with a standard deviation of 0.52, yielding a coefficient of variation of 10%. These values alone indicate that there is variation in syllable duration during conversational speech. But they do not indicate the extent or pattern of local variation within a single utterance and, indeed, given the method used to calculate articulation rate, the extent of variation is most certainly underestimated.

More specifically, articulation rate was computed in the following manner. For each speaker, each response to an interviewer's question was divided into stretches of speech that contained no pauses; these were called runs. The number of syllables in each run and its duration were determined. Next, successive runs of speech were grouped together, beginning at the onset of the first response, so that successive groups would each contain approximately 30 syllables. The articulation rate of each group ('utterance') was then computed by dividing the total number of syllables in the group by the cumulative duration of the runs comprising the group (excluding any pause time). With this procedure, then, a single value of articulation rate was based on approximately 30 syllables of speech; these 30 syllables typically involved more than one run from a single response and often runs from different responses. Note that this procedure is similar to that used by *Goldman-Eisler* [1956] in the respect that each data point represents the average across an often considerable stretch of speech. And it is precisely this type of measurement procedure that constitutes the measurement issue, the second major reason that it is premature to conclude that there is little variation in articulation rate during conversational speech. Namely, the data reported in the literature are, for the most part, based on averages across long stretches of speech, and thus do not reveal the extent of local variation in speaking rate within a given utterance of a given speaker.

The purpose of the present investigation was to begin to examine more closely the variation in articulation rate that occurs within a single utterance of a speaker during conversational speech. Specifically, we undertook a reanalysis of the English protocols of *Grosjean and Deschamps* [1975], concentrating on local variation in rate of articulation. In the next section of the paper we present the methodology and the results of this analysis. In the final section we consider the significance of our results for theories of speech perception and we discuss their implications for issues of speech planning.

The Reanalyses

Method

As indicated above, the original *Grosjean and Deschamps* [1975] analyses were based on interviews with 30 speakers. All interviews were taped from the BBC program 'The World at One'. Each speaker was being interviewed on a familiar topic, for example, the war in Northern Ireland, the adequacy of news coverage, or the arms race. Our analyses were based on the written protocols that were prepared for the original study. Each protocol consists of a written transcript of the speaker's answers to the interviewer's questions, with various linguistic and acoustic parameters marked on the protocol. The acoustic measurements were originally made from a visual display of the speech waveform on a chart recorder, and are accurate to the nearest 20 ms.

The data base from each interview thus consists of the speaker's responses to questions. Each response is divided into a series of runs. A run is defined as a stretch of speech that contains no pauses, with a pause defined as a silent interval of 250 ms or greater. The protocol indicates, for each response to a question, the number of syllables in each run of speech and the duration of the run. There is thus sufficient information in the protocols to compute the articulation rate of each individual run, i. e., articulation rate averaged over a single stretch of pause-free speech. As noted above, however, in the original analysis a single value of articulation rate was not computed over an individual run, but over as many runs as necessary to yield approximately 30 syllables of speech. For the new analysis, we computed articulation rate for each individual run of speech, within each response. Across all utterances the mean number of syllables per run was 11.68, with a standard deviation of 8.82. The rate of each run was computed in terms of milliseconds per syllable; this measure provides an index of the average duration of a syllable in a single run of speech. Of primary interest is the extent of variation in this measure of articulation rate for individual speakers.[1]

[1] In the original *Grosjean and Deschamps* [1975] study, for each speaker the first 15 groups of syllables (each approximately 30 syllables long) in the interview were included in the analysis. In nearly every case, this resulted in only part of the data from the final response being analyzed. Since it was important to base our analysis on only complete responses, we did not include the partially analyzed final responses in our data base. Thus, the two data bases are not absolutely identical.

Results

Summary data for the 30 speakers are presented in table I. Consider first the left section of the table, which provides statistics based on all responses of the speaker (ranging from one to five). The six measures are as follows: N is the total number of runs across all the responses; $\bar{x}$ and SD are the mean articulation rate (in ms/syllable) and standard deviation across all the runs; S is the articulation rate of the slowest run (i. e., longest average syllable duration), F is the articulation rate of the fastest run (i. e., shortest average syllable duration), and S–F is the difference in articulation rate between the slowest and fastest run. Clearly, across the interview, the articulation rate – in terms of average syllable duration in a stretch of pause-free speech – changed considerably for an individual speaker. The average standard deviation was 67 ms, which, for a mean of 216 ms, yields a coefficient of variation of 31%. Moreover, for all but 1 speaker (FLA) the difference in average syllable duration between the fastest and slowest run was over 100 ms, and for 20 of the 30 speakers it was 300 ms or greater; on average, the difference was 323 ms. Thus, across the interview, individual speakers effected large changes in the average duration of a syllable – on the order of hundreds of milliseconds.

Next consider the middle and right sections of the table. The middle section gives the data for the response of the subject during the interview that yielded the greatest difference in average syllable duration between the fastest and slowest run. On average, this difference was 309 ms; thus, even within a single response to a question, there can be considerable variation in rate. Finally, the right section provides data for the

Table I. Articulation rate data for 30 speakers based on all responses, the most variable response, and the least variable response. N = Total number of runs across the response(s); x̄ and SD = the mean articulation rate (in ms/syllable) and standard deviation across the runs; S = the articulation rate of the slowest run; F = the articulation rate of the fastest run; S−F = the difference in articulation rate between the slowest and fastest runs.

Speaker	Number of responses	All responses						Most variable response						Least variable response					
		N	x̄	SD	S	F	S−F	N	x̄	SD	S	F	S−F	N	x̄	SD	S	F	S−F
BAI	4	38	228	47	360	165	195	13	227	49	360	165	195	9	226	27	267	188	79
BLO	3	44	225	61	400	138	262	9	246	85	400	160	240	22	212	51	320	153	167
BRA	3	27	226	88	520	137	383	13	240	92	520	143	377	6	179	31	219	137	82
CAR	3	25	192	54	320	90	230	13	213	52	320	163	157	6	155	46	227	90	137
DUK	3	30	294	121	680	173	507	15	336	145	680	175	505	7	248	61	333	176	157
FIT	3	24	197	73	500	115	385	5	270	135	500	162	338	8	161	24	185	115	70
FLA	2	11	173	19	207	149	58	6	175	20	207	151	56	5	171	20	195	149	46
FOO	3	33	222	56	440	140	300	20	240	63	440	140	300	9	190	21	218	152	66
FRI	1	36	194	78	600	120	480	36	194	78	600	120	480	36	194	78	600	120	480
GAT	2	52	236	66	480	156	324	23	263	77	480	160	320	29	215	48	360	156	204
GOL	2	26	209	76	453	131	322	14	221	95	453	133	320	12	195	45	280	131	149
HAI	1	32	207	49	338	136	202	32	207	49	338	136	202	32	207	49	338	136	202
HAL	1	18	221	45	320	135	185	18	221	45	320	135	185	18	221	45	320	135	185
HUG	4	40	215	69	520	111	409	13	256	99	520	172	348	7	204	24	230	169	61
JAN	2	34	227	58	480	148	332	24	226	61	480	166	314	10	228	53	320	148	172
LEW	1	39	224	61	400	140	260	39	224	61	400	140	260	39	224	61	400	140	260
MAC	1	41	207	66	440	120	320	41	207	66	440	120	320	41	207	66	440	120	320
MIT	3	30	184	82	440	73	367	6	247	150	440	133	307	6	153	14	169	136	33
NEE	5	21	236	83	547	143	404	6	262	140	547	184	363	2	201	12	210	193	17
NEV	1	25	183	61	440	133	307	25	183	61	440	133	307	25	183	61	440	133	307
PUL	3	33	234	75	520	120	400	12	245	98	520	147	373	8	230	54	330	144	186
RAD	2	23	181	31	256	127	129	12	177	31	244	127	117	11	185	33	256	145	111
SAL	2	24	197	100	640	109	531	13	211	134	640	109	531	11	182	32	264	139	125
SAM	1	40	228	74	440	127	313	40	228	74	440	127	313	40	228	74	440	127	313
UND	4	33	194	68	440	120	320	11	211	85	440	133	307	5	185	26	220	151	69
VIC	3	31	204	54	370	115	255	15	206	60	370	115	255	3	200	26	227	175	52
WAR	2	30	214	49	360	73	287	15	219	52	320	73	247	15	209	47	360	160	200
WHI	2	39	240	67	460	92	368	25	240	74	460	92	368	14	240	56	400	171	229
WIN	2	35	259	89	560	150	410	26	270	98	560	150	410	9	226	43	307	168	139
WOO	1	22	228	86	560	110	450	22	228	86	560	110	450	22	228	86	560	110	450
Mean		31	216	67	450	126	323	19	230	80	448	139	309	16	203	44	314	146	169

response of the speaker that yielded the least change in average syllable duration between the fastest and slowest run. Even here, the variation in rate is considerable; on average, the difference between the fastest and slowest rate is 169 ms/syllable.

In addition to the extent of variation in rate, the pattern of variation across a response is of interest. Figure 1 (continuous lines) displays how speaking rate changes during the course of the interview for a speaker with an average degree of rate variation (GAT). For each of the two responses produced by the speaker, the articulation rate (ms/syllable) is plotted as a function of consecutive runs of speech. The most striking characteristic of the graph is that during a given response, articulation rate does not gradually increase or decrease, but changes course a number of times. This general pattern was typical of the 30 speakers in the corpus. Thus, the average duration of a syllable can change substantially not only over the course of an entire interview or even a single response during an interview, but between adjacent stretches of pause-free speech.

Finally, it is instructive to compare directly the pattern of results found in the current analysis with that reported by *Grosjean and Deschamps* [1975]. As noted above, those investigators reported a mean articulation rate, averaged across all speakers and groups (each consisting of roughly 30 syllables of speech), of 5.17 syllables/s with a standard deviation of 0.52, yielding a coefficient of variation of 10%. To obtain comparable statistics from our analysis, we converted the articulation rate of each run of each speaker from milliseconds per syllable to syllables per second. Averaging across speakers and runs, we obtained a mean rate

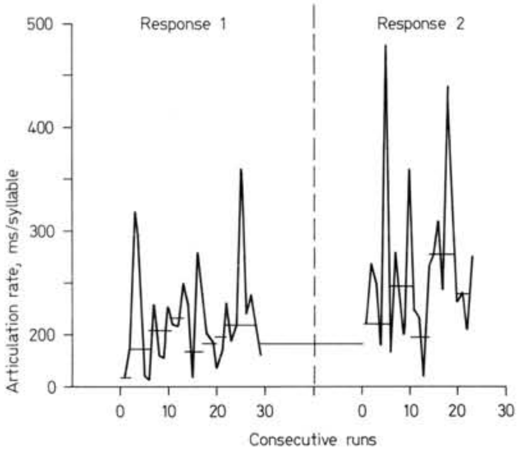

Fig. 1. Pattern of variation in articulation rate across two responses of a single speaker (GAT) in an interview. The continuous lines, which are based on the current analysis, display the articulation rate (in ms/syllable) as a function of consecutive runs of speech. The individual horizontal lines display the results of the analysis reported by *Grosjean and Deschamps* [1975]. Each horizontal line represents the mean articulation rate (in ms/syllable) across a group of syllables, each group being approximately 30 syllables in length.

of 4.97 with a standard deviation of 1.35, yielding a coefficient of variation of 27%.[2] Thus, as expected, the current analysis procedure reveals greater variability in rate than did the original. This can be seen graphically in figure 1. Superimposed on the graph of the current analysis (continu-

[2] In order to ensure that the difference between studies was not due to the fact that slightly different data bases were used (see footnote 1), we recomputed the statistics for the individual runs on the entire data base used by *Grosjean and Deschamps* [1975], that is to say, on the incomplete final responses as well as the complete responses. The mean articulation rate (syllables/s) across speakers and runs was 4.96 and the standard deviation was 1.34, values nearly identical to those based on complete responses only (4.97 and 1.35, respectively).

ous lines) is a display of the results from the original study (individual horizontal lines). Each horizontal line represents the mean articulation rate across a single group of (approximately 30) syllables. With a single exception, every group involves more than one run and, in addition, one group includes runs from both the first and second responses. It is readily apparent from the superimposed graphs that the averaging technique used in the original analysis greatly underestimated the variation in rate.

Discussion

The most important finding of the present investigation is that during conversational speech the variation in average syllable duration for a given speaker is substantial, typically on the order of hundreds of milliseconds, and given that our analysis was based on the average syllable duration over a stretch of speech (i.e., a run of pause-free speech), and not on individual syllable durations, the results provide at best a conservative estimate of the extent of variability in articulation rate.

Our findings raise questions about the results obtained by earlier researchers [*Goldman-Eisler*, 1968; *Grosjean and Deschamps*, 1975; *Malécot* et al., 1972] as well as about the interpretations that accompany them. It would appear that too much emphasis has been placed on the rather 'low' variation in articulation rate as compared to the rather 'high' variation in pause rate (the magnitude of variation issue), and this has led researchers to conclude that articulation rate per se does not vary substantially either between or within speakers. In addition, articulation rate typically has been measured over large stretches of speech, such that the local variation characterized by the peaks and troughs that can be seen in figure 1 is neutralized (the measurement issue). Our reanalysis clearly shows that articulation rate varies quite considerably within and across speakers, even when the unit of analysis is an entire run of pause-free speech, and not an individual syllable. This finding raises interesting issues concerning both speech perception and speech production.

As for speech perception, the variation we have found is significant with respect to the durational properties that specify many segmental distinctions. Consider, as an example, the segmental distinction between syllable-initial stop consonant /b/ and semivowel /w/. A major distinguishing characteristic of /b/ and /w/ is the abruptness of the consonantal onset, with the onset for /b/ being more abrupt than that for /w/ [*Dalston*, 1975; *Fant*, 1960; *Fischer-Jørgensen*, 1954]. One aspect of abruptness that is perceptually relevant is the duration of the initial formant transitions: Numerous experiments have shown that syllables with short initial transitions are perceived as beginning with a stop consonant whereas those with longer transitions are perceived as beginning with a semivowel [*Cooper* et al., 1976; *Liberman* et al., 1956; *Schwab* et al., 1981].

In a recent investigation, *Miller and Baer* [1983] examined how syllable duration and transition duration for the initial consonants in the syllables /ba/ and /wa/ change as a function of speaking rate. They asked speakers to produce these syllables at rates ranging from very fast to very slow. For both /ba/ and /wa/, the speakers produced syllables ranging from under 100 ms to over 600 ms, a range that is comparable to the range of average syllable durations produced by speakers in the current study (table I). Of particular interest was how transition duration changed as a function of syllable duration. As /ba/ became longer, transition duration remained relatively constant at approximately 40–50 ms; however, as the duration of /wa/ increased, so too did the transition duration, from a value of approximately 50 ms for the shortest syllables to one of approximately 140 ms for the longest syllables. With this pattern of results, the criterion transition duration that would most effectively differentiate /ba/ from /wa/ is not constant, but rather increases as syllable duration increases. And, indeed, perceptual studies have shown that listeners do process transition duration in such a rate-dependent manner: As the duration of the syllable increases, an increasingly longer transition duration is required to perceive /wa/ as opposed to /ba/ [*Miller and Liberman,* 1979]. Interestingly, the largest effects of syllable duration on the perception of /b/ and /w/ have been found across syllable durations ranging from approximately 80 to 200 ms – a range that is covered by speakers in natural conversation, as table I indicates. Of course the /ba/ -/wa/ data are based on individual syllable durations whereas the current analysis considers only average syllable duration across a stretch of speech. Nevertheless, the comparison suggests that the type of rate-dependent processing found in the perceptual experiments with isolated syllables may well be critical to perception of spontaneous speech, and it sets the stage for more fine-grained acoustic analysis of individual syllables in conversational speech.

Our findings are not only relevant to theories of segmental perception, as discussed above, but have interesting implications for models of speech planning. One question of interest concerns the factors that lead to the variability we have observed in articulation rate. Researchers such as *Goldman-Eisler* [1968] have isolated 'macro variables' that influence rate, that is, those variables that have an effect on the rate of the global response. Among these we find the emotional state of the speaker, the type of speaking situation, the familiarity of the material being discussed, and the like. But because researchers were unaware of the variation in local articulation rate, they have not studied the 'micro variables', that is, those variables that have local but also quite large effects on change of rate (as we have seen). Future research will want to isolate the many factors that play such a role, for example, lexical access difficulties, syntactic construction delays, semantic planning problems, and so forth, and to account for them in models of planning and execution of speech.

A second question of interest concerns the domains over which a change in articulation rate can occur. More specifically, can articulation rate be modified within any domain – phonological, morphological, prosodic, syntactic, discourse – or must it sometimes await a particular domain boundary? A careful acoustic analysis of the speech stream in relation to the many potential domains would be able to provide an answer to this question. It may be, for example, that the production system can only make rate changes at precise points: between syllables but never within, between phonological words but not within, between phrases (syntactic or prosodic) but not within, etc. Identifying these points will not only contribute to our understanding of the underlying mechanisms controlling rate of speech, but will also provide basic information about the on-line planning and execution of language.

In summary, our analyses have shown that in a conversational setting, talkers effect substantial changes in the rate at which they produce speech, such that average syllable durations may span a range of hundreds of milliseconds. The task for future research is to explicate the operations in the planning and execution of language that underlie these changes, and to specify the manner in which listeners are able to accommodate for these changes in the course of speech perception.

Acknowledgements

This research was supported by NIH grant NS 14394 and RCDA NS 00661 to J.L.M., by NSF grant BNS 8404565 to F.G., and by NIH BRSG RR 17043 to Northeastern University. We express our appreciation to *James Gee* for his insightful comments throughout the course of the study.

References

Cooper, W. E.; Ebert, R. R.; Cole, R. A.: Perceptual analysis of stop consonants and glides. J. exp. Psychol. human Perception Performance *2:* 92–104 (1976).

Dalston, R. M.: Acoustic characteristics of English /w,r,l/ spoken correctly by young children and adults. J. acoust. Soc. Am. *57:* 462–469 (1975).

Fant, G.: Acoustic theory of speech production (Mouton, The Hague 1960).

Fischer-Jørgensen, E.: Acoustic analysis of stop consonants. Miscellanea phonetica *ii:* 42–59 (1954).

Goldman-Eisler, F.: The determinants of the rate of speech output and their mutual relations. J. psychosom. Res. *1:* 137–143 (1956).

Goldman-Eisler, F.: Psycholinguistics: experiments in spontaneous speech (Academic Press, London 1968).

Grosjean, F.: Temporal variables within and between languages; in Dechert, Raupach, Towards a cross-linguistic assessment of speech production (Lang, Frankfurt am Main 1980).

Grosjean, F.; Deschamps, A.: Analyse des variables temporelles du français spontané. II. Comparison du français oral dans la description avec l'anglais (description) et avec le français (interview radiophonique). Phonetica *28:* 191–226 (1973).

Grosjean, F.; Deschamps, A.: Analyse contrastive des variables temporelles de l'anglais et du français: vitesse de parole et variables composantes, phénomènes d'hésitation. Phonetica, *31:* 144–184 (1975).

Jusczyk, P. W.: A review of speech perception research; in Kaufman, Thomas, Boff, Handbook of perception and performance (Wiley, New York, in press, 1984).

Liberman, A. M.; Cooper, F. S.; Shankweiler, D. P.; Studdert-Kennedy, M.: Perception of the speech code. Psychol. Rev. *74:* 431–461 (1967).

Liberman, A. M.; Delattre, P. C.; Gerstman, L. J.; Cooper, F. S.: Tempo of frequency change as a cue for distinguishing classes of speech sounds. J. exp. Psychol. *52:* 127–137 (1956).

Malécot, A.; Johnston, R.; Kizziar, P.-A.: Syllabic rate and utterance length in French. Phonetica *26:* 235–251 (1972).

Miller, J. L.: Effects of speaking rate on segmental distinctions; in Eimas, Miller, Perspectives on the study of speech (Erlbaum, Hillsdale 1981).

Miller, J. L.; Baer, T.: Some effects of speaking rate on the production of /b/ and /w/. J. acoust. Soc. Am. *73:* 1751–1755 (1983).

Miller, J. L.; Liberman, A. M.: Some effects of later-occurring information on the perception of stop consonant and semivowel. Perception Psychophysics *25:* 457–465 (1979).

Nooteboom, S. G.; Brokx, J. P. L.; Rooij, J. J. de: Contributions of prosody to speech perception; in Levelt, Flores d'Arcais, Studies in the perception of language (Wiley & Sons, Chichester 1978).

Pisoni, D. B.: Speech perception; in Estes, Handbook of learning and cognitive processes, vol. 6 (Erlbaum, Hillsdale 1978).

Port, R. F.: Linguistic timing factors in combination. J. acoust. Soc. Am. *69:* 262–274 (1981).

Schwab, E. C.; Sawusch, J. R.; Nusbaum, H. C.: The role of second-formant transitions in the stop-semivowel distinction. Perception Psychophysics *29:* 121–128 (1981).

Stevens, K. N.; Blumstein, S. E.: The search for invariant acoustic correlates of phonetic features; in Eimas, Miller, Perspectives on the study of speech (Erlbaum, Hillsdale 1981).

Summerfield, A. Q.: Aerodynamics versus mechanics in the control of voicing onset in consonant-vowel syllables; in Speech perception, No. 4 (Department of Psychology, Queen's University of Belfast 1975).

Summerfield, Q.: On articulatory rate and perceptual constancy in phonetic perception. J. exp. Psychol. human Perception Performance *7:* 1074–1095 (1981).

Tuller, B.; Harris, K. S.; Kelso, J. A. S.: Stress and rate: differential transformations of articulation. J. acoust. Soc. Am. *71:* 1534–1543 (1982).

Received: September 17, 1984
Accepted: October 4, 1984

Dr. Joanne L. Miller,
Department of Psychology,
Northeastern University,
Boston, MA 02115 (USA)