THE LINGUISTIC BASIS OF ARTICULATION RATE VARIATION IN CZECH

BY

JANA DANKOVIČOVÁ

OXFORD UNIVERSITY

ST. HUGH'S COLLEGE

TRINITY TERM 1998

A thesis submitted to the Committee for Comparative Philology and General Linguistics in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Oxford University
Phonetics Laboratory

THE LINGUISTIC BASIS OF ARTICULATION RATE VARIATION IN CZECH

JANA DANKOVIČOVÁ

St. Hugh's College, University of Oxford
D. Phil. dissertation, Trinity Term 1998

ABSTRACT

The thesis presents an experimental phonetic investigation of variation in articulation rate in Czech. It reports on two studies. In the first study, I investigated whether variation in articulation rate has a specific domain in elicited spontaneous Czech speech. Three types of unit were examined as candidate domains: the interpause stretch, the intonation phrase and the clause. Articulation rate within these units was measured for each (phonological) word. The analysis showed that articulation rate was not constant in any of these candidate domains. However, a regular pattern was found within the intonation phrase, consisting of a slowing down (rallentando) with the first or the second word being the fastest in the majority of intonation phrases. Although phrase-final lengthening contributes to the decrease at the end of the intonation phrase, other factors contribute to the overall pattern. As one possible determinant of articulation rate, the word size was considered. A statistically significant tendency was found for the articulation rate of the word to increase with its size, expressed in the number of syllables. The hypothesis that the interpause stretch and/or clause might constitute hierarchically higher domains of variation in articulation rate was not confirmed. However, some evidence was found for the interpause stretch as a superordinate domain of phrase-final lengthening.

In the second study, a tendency of the articulation rate to decrease throughout the intonation phrase was tested statistically, using a considerably larger database, including three different tasks: reading a story, paraphrasing the story and elicited spontaneous speech. The rallentando pattern was confirmed for all three tasks. Moreover, a high correlation between speakers was obtained for the articulation rate of individual words when speakers were reading the same text. In the investigation of possible determinants of the word's articulation rate, the same tendency as in Study 1 was observed for word size in syllables. In addition, I found a tendency for phonological words containing only function words to be significantly faster than phonological words containing only content words. The finding of a significant effect of word size and word class on the articulation rate of words is in accord with the results reported in the literature for other languages.

Finally, a mathematical model was designed to estimate and predict articulation rate on the basis of three variables: position of the word within the intonation phrase, its size in syllables and word class.

AKNOWLEDGEMENTS

I should like to thank first of all my supervisor Dr. John Coleman, for his constant support and invaluable help and discussions throughout my work on this thesis and for keeping my feet and ideas on the ground while reading all my successive drafts.

Sincere thanks are also due to Dr. Francis Nolan for his always-ready intellectual and moral support during my work on this project, to Prof. William Barry and Prof. Janet Pierrehumbert for a number of useful suggestions, and to Dr. Ian Watson both for some very useful discussions in the initial stages and for his last-minute emergency advice on statistics.

I am very grateful to the reviewers of my paper published in the Journal of Phonetics, Professors Janet Fletcher, Ilse Lehiste, Patricia Speetor-Beddor. I found their advice most valuable for extending this work into a D. Phil. thesis.

For crucial, continuous, help with statistical analysis I am particularly indebted to Dr. Mario Cortina-Borja. The amount of time and patience he put into our meetings was most valued. I should also like to thank Dr. Paul Griffiths for his statistical advice on few occasions and to Mr. Peter Higginbotham, from Oxford University Computing Services, for recovering my files from the diskette when they seemed lost.

With respect to general technical support, I am most grateful to Andrew Slater, Senior IT Support Officer, who gave me always prompt and first-class assistance when problems arose with computers.

My thanks also go to Charles University, Prague, my home university, which supported my application to come to Oxford to study linguistics. I am very grateful to the members of the Institute of Phonetics, who allowed me to use part of their research material and their recording studio when collecting the rest of my material. They were also the first people who aroused my interest in phonetics. In this respect, I should like to express my thanks especially to Prof. Zdena Palková and Prof. Marie Dohalská. My thanks to Prof. Palková also for lending me her ears and skills in dealing with some difficult aspects of the data and for her comments on earlier stages of this work. I am also grateful to Prof. Přemysl Janota and Prof. Miroslav Ptáček for their comments and suggestions when starting this project.

This thesis could have not been produced without the support of St. Hugh's College, Oxford, which awarded me a Rawnsley Studentship, thus making it possible to pursue my studies in Oxford. In this respect, I am also grateful to the Soros Foundation/FCO, who provided the funding for my first year in Oxford, and the Dulverton Trust for further financial support at a later stage. For generous financial support to attend various conferences I would like to thank St. Hugh's College and the Phonetics Laboratory.

Last but not least, I should like to thank all the people in the Phonetics Laboratory, Oxford, for providing such a pleasant atmosphere throughout my research. My thanks also go to all my friends for keeping my social life afloat. Finally, I thank Daniel warmly for proof-reading my final draft, restraining my attempts to reform the English language, and for making me eat something more than KitKat during the last few days.

CONTENTS

Abstract i
Acknowledgements ii
List of Tables viii
List of Figures x

Chapter 1 INTRODUCTION

1.1 Preliminary remarks 1
1.2 The main issues in the research on articulation rate 3
1.3 What answers does the literature give? 7
1.3.1 Problems with methodology 7
1.3.2 Articulation rate variation across speakers and within the speech of an individual 9
1.3.3 Speech rate, articulation rate and pauses 13
1.3.4 Speech rate and articulation rate variation, and independent variables 14
1.3.5 Other determining factors 19
1.3.6 Do languages differ with respect to articulation rate? 26
1.4 Aims and outline of the thesis 30

Chapter 2 STUDY 1: EXPERIMENTAL DESIGN

2.1 Introduction 34
2.2 Material 37
2.3 Subjects 37
2.4 Recording procedure 38
2.5 Methodology for the analysis of articulation rate variation 39
2.5.1 Candidate domains: definitions 39
2.5.1.1 Interpause stretch 39
2.5.1.2 Clause 40
2.5.1.3 Intonation phrase 40
2.5.2 Candidate domains: justification for their choice 42

2.5.3 The unit for the measurement 43
2.5.4 The phonological word in Czech 45
2.5.5 Measurement technique 46
2.5.6 The unit for expressing articulation rate 48
2.6 Determinants of articulation rate variation 50
2.6.1 Position of the phonological word within the domain 50
2.6.2 Size of the phonological word 51
2.7 Statistical processing 51

Chapter 3 STUDY 1: RESULTS

3.1 Frequency distribution of intonation phrases of different sizes 53
3.2 Is articulation rate constant within intonation phrases? 54
3.3 Patterning of articulation rate variation within intonation phrases 56
3.4 Internal structure of articulation rate variation within the intonation phrase 60
3.4.1 Assumptions for GLM analysis 62
3.4.2 Effect of the position of the word within the intonation phrase and its size on its articulation rate: An overall model 63
3.4.3 Patterns ABC and BAC 68
3.4.4 Word size and articulation rate 72
3.5 Internal structure of articulation rate variation within the interpause stretch and the clause 79
3.5.1 Interpause stretches 80
3.5.2 Clauses 82
3.6 Phrase-final lengthening 85
3.6.1 Interpause stretches 85
3.6.2 Clauses 86
3.7 Conclusion 87

Chapter 4 STUDY 2: EXPERIMENTAL DESIGN

4.1 Introduction 90
4.2 Material 94
4.3 Subjects 96
4.4 Recording procedure 97
4.5 Methodology for the analysis of articulation rate variation 98
4.6 Measurement technique 99
4.7 Determinants of articulation rate variation 100
4.7.1 Position of the phonological word within the intonation phrase 100
4.7.2 Size of the phonological word 100
4.7.3 Function vs. content word 101
4.8 Statistical processing 103

Chapter 5 STUDY 2: ELEMENTARY RESULTS

5.1 Frequency distribution of intonation phrases of different sizes 105
5.2 Inter-speaker correlations of articulation rate of phonological words (reading) 107
5.3 Patterning of articulation rate variation within the intonation phrase 112
5.3.1 Two-word intonation phrases 113
5.3.2 Three-word intonation phrases 113
5.3.3 Four-word intonation phrases 115

Chapter 6 STUDY 2: RESULTS OF THE STATISTICAL ANALYSIS

6.1 Assumptions for the GLM analysis 119
6.2 Overall model 120
6.2.1 Model with the main factors, 2-way and 3-way interactions 121
6.2.2 Results of the main analysis 122

6.2.2.1 Position 125
6.2.2.2 Word size 130
6.2.2.3 Word class 133
6.2.2.4 Relationships between position, word size and word class 135
6.2.2.5 Speaker 138
6.2.2.6 Task 140
6.2.2.7 Patterns ABC and BAC 142
6.3 Conclusion 149

Chapter 7 MATHEMATICAL MODELLING OF ARTICULATION RATE

7.1 Introduction 153
7.2 Model 1 157
7.3 Model 2 160
7.4 Model 3 160
7.5 Model 4 163
7.6 Model 5 164
7.7 Model 6 166
7.8 Conclusion 167

Chapter 8 SUMMARY AND DISCUSSION

8.1 Intonation phrase as a domain of articulation rate variation 169
8.2 Factors affecting articulation rate of phonological words 177
8.3 Points about methodology and more on the intonation phrase 182
8.4 Applications 185

Appendix 1 188
Appendix 2 200
Appendix 3 222
Appendix 4 235
Appendix 5 238
Appendix 6 241
Appendix 7 249

Appendix 8 269
Appendix 9 278
Appendix 10 293
References 318

LIST OF TABLES

3.1	Frequency distribution of intonation phrases according to their size	53
3.2	Ordinal patterns for 2-word intonation phrases	57
3.3	Ordinal patterns for 3-word intonation phrases	57
3.4	Ordinal patterns for 4-word intonation phrases	58
3.5	Mean articulation rates [sylls/s] (summed across speakers) and standard deviations for words in different positions within the intonation phrase, and the results of the Tukey test ($p < 0.05$) on which positions involved significantly different means	66
3.6	Mean articulation rates [sylls/s] and standard deviations for words in different positions within the intonation phrase: the results for individual speakers in 3-word phrases. The results of the Tukey tests ($p < 0.05$) on which positions involved significantly different means	67
3.7	Patterns ABC and BAC: Mean articulation rates [sylls/s] and standard deviations for words in different positions within the intonation phrase, and the results of the Tukey test on which positions involved significantly different means ($p < 0.05$)	69
3.8	The levels of significance from one-way ANOVA for the factors of number of syllables and number of segments, taken separately, and with articulation rate as the dependent variable	73
3.9	Average number of segments per syllable as a function of number of syllables in a word	76
3.10	*t* values for tests of differences between articulation rates (AR) of final words in non-final intonation phrases within the interpause stretch and final words in the final intonation phrase within the interpause stretch	85
3.11	*t* values for tests of differences between articulation rates (AR) of final words in non-final intonation phrases within the clause and final words in the final intonation phrase within the clause	86
4.1	The duration of speech samples used in the study	95

Oxford University
Phonetics Laboratory

5.1 The frequency distribution of intonation phrases in categories of 1-word, 2-word, 3-word, 4-word and 5-or-more word intonation phrases 106

5.2 Inter-speaker correlations of the articulation rate of phonological words in reading (Pearson's correlation coefficient *r*) 110

5.3 Ordinal patterns for 2-word intonation phrases 113

5.4 Ordinal patterns for 3-word intonation phrases 114

5.5 Ordinal patterns for 4-word intonation phrases 116

6.1 Results of the GLM analysis for 2-word, 3-word and 4-word intonation phrases 124

6.2 Mean articulation rates [sylls/s] and standard deviations (SD) summed across speakers and tasks, for words in different positions within the intonation phrase, and the results of the Tukey test ($p < 0.05$) on which positions involved significantly different means 126

6.3 Mean articulation rates [sylls/s] and standard deviations (SD) for words in different positions within the intonation phrase: the results for individual speakers in 3-word phrases (tasks combined). The results of the Tukey tests ($p < 0.05$) on which positions involved significantly different means 128

6.4 Categories of words with different size in syllables which were significantly different from each other. (The results of the Tukey test) 132

6.5 Word class categories which were significantly different from each other. (The results of the Tukey test) 134

6.6 Speakers who were significantly different from each other. (The results of the Tukey test) 140

6.7 Patterns ABC and BAC: Mean articulation rates [sylls/s] and standard deviations for words in different positions within the intonation phrase, and the results of the Tukey test on which positions involved significantly different means ($p < 0.05$) 144

7.1 Estimates for goodness of fit in Models 1-6 167

LIST OF FIGURES

2.1	A waveform and a f_0 track of three intonation phrases from speaker MS, with vertical lines indicating the measurement points	48
3.1	Frequency distribution of intonation phrases according to their internal articulation rate variation	55
3.2	Distribution of phonological words in different positions within the intonation phrase according to their size in syllables	71
3.3	Correlation between the number of segments and the number of syllables in the word	74
3.4	Correlation between the articulation rate and the number of segments in the word	76
3.5	Correlation between the articulation rate and the number of syllables in the word	77
3.6	Correlation between the articulation rate and the number of segments per syllable in the word	78
3.7	Articulation rate within interpause stretches (Speaker OL)	81
3.8	Articulation rate within clauses (Speaker OL)	83
3.9	Articulation rate within clauses - words 109-170 (Speaker JP)	84
5.1	Articulation rate variation in reading: individual speakers	108
5.2	Frequency distribution of phonological words according to their degree of variation across different speakers	110
6.1	Boxplot for articulation rate in different positions within the intonation phrase	127
6.2	Boxplot for articulation rate in different positions within the intonation phrase: individual speakers	129
6.3	Boxplot for articulation rate in words with different number of syllables	132

6.4	Boxplot for articulation rate in words of different word class	134
6.5	Distribution of categories of word class for words consisting of different number of syllables	137
6.6	Distribution of word class categories in different positions within the intonation phrase	138
6.7	Boxplot for articulation rate in different speakers	139
6.8	Boxplot for articulation rate in different tasks	141
6.9	Distribution of words according to their size in syllables for different positions within the intonation phrase	145
6.10	Distribution of words according to their word class category for different positions within the intonation phrase	147
7.1	Scatterplot of the articulation rate as a function of size in syllables	162

CHAPTER 1

INTRODUCTION

1.1 PRELIMINARY REMARKS

The term 'articulation rate' is traditionally used as a measure for the rate of speaking[1] in which all pauses are excluded from the calculation. It contrasts with the other way of measuring rate of speaking - 'speech rate', which includes pauses as well as articulation rate. Because speech rate is a complex measure, it is not as informative as examining articulation rate, pause duration and pause frequency separately, since different speakers of the same language or speakers of different languages may use various strategies for combining speaking and pausing. Thus, studies dealing with detailed timing aspects of speech either examine both articulation and speech rates, or employ the measure of articulation rate only. When the concern is modelling timing in speech, dealing separately with articulation rate, on the one hand, and pause duration and frequency, on the other, is inevitable. An important part of such research is investigating how articulation rate changes over time, in other words, what the variation in articulation rate is. (In some studies on this topic, e.g., Crystal & House, 1990, the term 'variation' is used synonymously with 'variability'.)

In a number of areas of phonetics research, particularly in modelling speech production and perception, one of the most problematic issues is the extent and nature of variation in the signal. Part of this variation is temporal, both at the segmental and prosodic levels. The research on segmental duration has been more extensive than the quantitative research on other aspects of timing. Most work on

[1] I use this term as an umbrella term covering both speech rate and articulation rate in places where the distinction is not necessary.

segmental duration has been done in relation to text-to-speech synthesis (for a summary of factors affecting segmental duration, see Klatt, 1976; van Santen, 1992). Duration of segments is given partly by their inherent duration, however, actual segmental duration is affected by a number of other factors (e.g., segmental environment and degree of stress). It has been shown repeatedly that articulation rate, as a prosodic factor, is one of the crucial determinants of actual segmental duration (e.g., Lindblom, 1963; Kozhevnikov & Chistovitch, 1965; Gay, 1981; Miller, 1981; den Os, 1988; Kaiki *et al.*, 1990; Carlson, 1991).

Interest in modelling prosody has been growing in recent years, especially for the purpose of text-to-speech synthesis. As Silverman *et al.* (1992) claim:

> "Prosody is central to the interface between speech and natural language processing technologies. It not only accounts for much of the variability in speech signals, but also conveys much of the information that is necessary for recovering the intended meaning of an utterance - information which is unavailable in orthographic transcriptions. Consequently an understanding of prosody - how it related both to the acoustic speech signal and to text and discourse structure - is crucial as speech understanding and synthesis technologies progress towards the developments of complete spoken language systems that accomplish complex real-world tasks." (Silverman *et al.*, 1992: 867)

An extensive body of research has been accomplished particularly on intonation modelling (e.g., Vaissière; 1971; Olive, 1974; O'Shaughnessy, 1976; Pierrehumbert, 1981). Models of various aspects of prosodic timing, namely pauses and overall timing structure, have not been explored to such an extent. Some phenomena, such as utterance-final syllable lengthening, have been incorporated in text-to-speech synthesis for some time, while others, such as control of pause duration, articulation rate and speech rhythm, are not yet commonly used. Several reasonably successful systems, which do allow for pauses and rhythmical constraints, have been proposed (for a review, see Zellner, 1994). However, variation in articulation rate is to a great extent still ignored in this area, no doubt due largely to the fact that articulation rate as such, and its

variation in particular, are not well-understood phenomena. Yet it is believed that introducing the factor of variation in articulation rate would contribute to improving both the naturalness of the signal (e.g., Carlson *et al.*, 1979; Eefting, 1990) and the accuracy of automatic speech recognition performance (e.g., O'Shaughnessy, 1995).

1.2 THE MAIN ISSUES IN THE RESEARCH ON ARTICULATION RATE

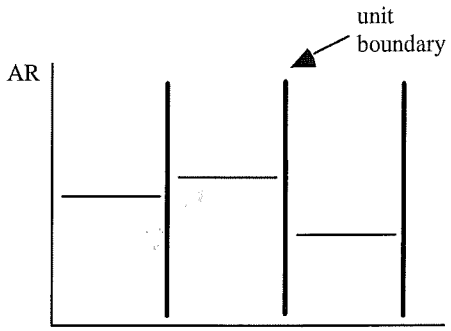
There are a number of questions one could ask when investigating articulation rate. I shall mention those which, in my opinion, are the most crucial when approaching the issue from the perspective of speech production. The discussion of the literature, which will follow in section 1.3, will provide a survey of the answers available so far.

- One of the most important questions is **how to measure articulation rate**. This might seem rather a trivial thing to ask but it is far from obvious what the answer is. A decision needs to be made especially about what the minimum unit on which to base the measurements should be. The decision will depend on the aim of an inquiry and a theory of articulation rate functioning. Another decision concerns the measure for expressing articulation rate; should we use the number of syllables per unit of time, the number of words or the number of sounds/phonemes? I shall discuss later (section 1.3.6) how the choice of articulation rate measure can influence the results, particularly, when comparing languages with a different syllable structure and a different average word length.

- Considering the situation within a single language, **does articulation rate vary across speakers and within the speech of an individual? If it does, to what extent?** A related, and problematic, issue is to define 'similar' and 'different' or 'considerably' different rate.

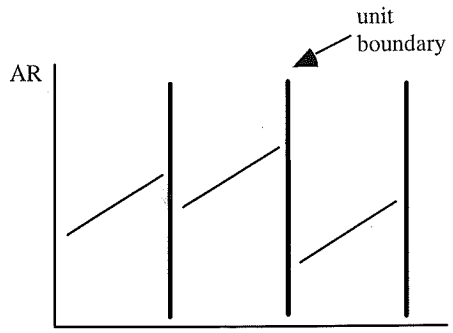
- If within-speaker variation in articulation rate can be found, two aspects should be distinguished: listener-oriented variation, which I shall call 'discourse variation', and 'speech-inherent variation'. **If evidence for inherent articulation rate variation within speakers does exist, is this variation arbitrary, or does the rate vary in a systematic way, governed by certain rules?**

- An important question, related to the previous one, is **does articulation rate have a domain of variation, a unit within which it is organised?** Two possibilities exist of how the rate can be organised within the unit. First, the rate may be constant within the unit and any change in the rate is restricted to the unit boundaries. This possibility is represented in the idealisation below.[2]

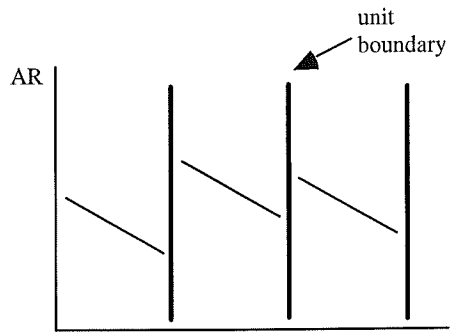


The alternative is that articulation rate may follow a certain pattern within the unit, a pattern whose shape would be recurrent in all units. For instance, the rate may progressively rise throughout the unit:

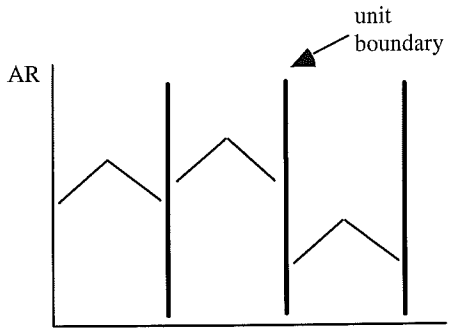
[2] The relationship between individual units in terms of their articulation rate as depicted (higher/lower) is completely arbitrary and irrelevant since what concerns us here is what is happening *within* the unit, not *between* the units.



or it may progressively decrease throughout the unit:



or it may manifest a more complex pattern, for instance, a rise followed by a fall, as in the picture below:



Although the above basic possibilities for organizing articulation rate variation within the unit (constant vs. pattern) were presented as alternatives, this is not to say that they cannot co-occur. An interesting issue to investigate would then be

whether the existence and distribution of these cases is related in some way to the linguistic structure of the units or some other factors (discourse, emotional, etc.).

- Suppose articulation rate proves to vary systematically. The question arises then: **what causes articulation rate to vary?** In other words, what are the main determinants of rate variation? There are a number of factors which are possible 'candidates'. First of all, independent variables, namely, age, sex, education, social class, emotional state, task or topic the speaker is engaged with, degree of preparation, behaviour of and familiarity with other people involved in the conversation. What would be the domain over which such factors operate is a further question. We should also ask how articulation rate is related to motor control, cognitive activities or whether it varies for the sake of the listener - for instance, to mark boundaries of some linguistic (e.g., syntactic, semantic, discourse) units and thus facilitate the process of decoding. Furthermore, does it relate to the linguistic structure of units? Is it a mixture of several factors of this kind? If so, a further question arises: how to separate these factors in the analysis to make it possible to estimate their relative contribution?

- **Given that both articulation rate and pauses are timing phenomena, could they be interrelated in some way in speech?**

- **If there are any consistent answers to any of the above questions, do they hold in other languages?** Especially, if any systematic behaviour of articulation rate is found in one language and it seems to be partly determined by the linguistic structure of this language, it is possible that another language behaves differently with respect to rate variation.

Answering the major questions outlined above would lead to establishing a theory of articulation rate, to the present day absent. Once we have a theory, we can test it from other perspectives, for instance, a developmental perspective, or a perspective of second-language users.

1.3 WHAT ANSWERS DOES THE LITERATURE GIVE?

Study of the literature suggests that of the questions outlined above, only few have been addressed. In this section, I shall describe the research concerned with these issues. The discussion of the literature aims to be quite detailed since a comprehensive survey is not, to my knowledge, available. However, one should bear in mind that the survey is by no means complete.

1.3.1 Problems with methodology

Rather strikingly, in most studies on speech and/or articulation rate the authors do not discuss why they adopt a particular methodology and not some other. The decision is simply stated as an obvious fact. This is true both about the unit across which articulation rate is measured and about the measure in which rate is expressed. I shall comment here only on the former issue and leave the latter one for sections 1.3.6 and 2.5.6.

Apart from failing to justify the unit used, some studies do not even provide a sufficient definition of this unit (e.g., Fant *et al.*, 1991a). Furthermore, the units in different studies vary in length, and this makes comparisons of results difficult. Another drawback is that the larger the unit employed, the more variation is averaged out. I shall comment on this problem in greater detail in section 1.3.2.

To my knowledge, there are four different units used in the literature in the past: (i) speech stretches between two successive questions of the interviewer (Goldman-Eisler, 1968), (ii) 30-syllable long stretches of speech (Grosjean & Deschamps, 1975), (iii) clause or sentence (e.g., Fant *et al.*, 1991a), and (iv) interpause stretches, i.e., stretches of speech between two successive pauses (e.g., Malécot *et al.*, 1972; Butcher, 1981; Miller *et al.*, 1984).

The first unit is a rather artificially defined discourse unit, normally too long to expose the variation in articulation rate. Furthermore, it is too variable in length. The second unit, 30-syllable long stretches, is an arbitrary unit, but it seems objective as it uses stretches of speech of equal length. The problem with this unit is that if articulation rate is structured in some way and if such structuring is related to some linguistic unit, this fact would be impossible to observe. The third unit, the clause or the sentence, is a relatively short linguistic unit. Hence, its use for measuring articulation rate variation is better justified than using the previous two units. The fourth unit, the interpause stretch, is probably the most often used nowadays.[3] One problem is that it is not clear what kind of unit the interpause stretch is. It may be a discourse unit or a planning unit; it is difficult to decide as pauses may fulfil several functions at the same time. The other difficulty is that the definition of the interpause stretch is not unified since authors differ in their definition of pause and the minimum pause duration. A few remarks about these issues are in place.

The term 'pause' is typically used in the literature for two kinds of breaks in speech that are assumed to play some role in the encoding process. The first group involves an instrumentally detectable drop in the intensity of the speech signal over a given period of time. These are called 'silent' pauses. The other group in the traditional division is called 'filled' pauses and these include 'er, ehm, ah' phenomena. In the research on speech and articulation rates, silent and filled pauses are either grouped together or only silent pauses are taken into account.

The minimum pause duration is defined as the cut-off point between pause as an encoding (cognitive) phenomenon, and the so called 'articulatory pause' which is usually defined as a necessary break in speech resulting from articulatory activity,

[3] Here we find an exception in the literature with respect to explicitly stating reasons for employing the unit chosen; Miler *et al.* (1984) justify using interpause stretches in their reanalysis of Grosjean and Deschamps' (1975) material to demonstrate that shorter units of speech show more variation than longer ones. However, they still do not explain why they chose the interpause stretch and not some other unit.

i.e., a period of silence which is not assumed to be used for speech planning (this is a case with stop consonant closures).

A decision as to what will be the minimum cut-off point in the analysis is crucial since calculating articulation rate involves subtracting the total duration of all pauses. Too high a cut-off point will make the articulation rate lower than it would be if minimum pause duration were smaller. There is no clear answer to the question of what the minimum pause duration should be. The overview of data on storytelling by Kowal *et al.* (1983) shows that the most usual value accepted is between 0.2 and 0.3 seconds, usually 0.25 seconds. This figure comes from Goldman-Eisler's work (1961a, 1968) and for a long time it was automatically adopted by a number of investigators. However, the tenability of such a high cut-off point is challenged in Heike *et al.* (1983) where Goldman-Eisler's decision is criticized as being rather arbitrary and to a certain extent due to instrumental limitations. In their work, based on phonotactic and psycholinguistic research, Heike *et al.* demonstrate that pauses shorter than 0.25 seconds are not primarily articulatory in origin and that pauses between 0.13 and 0.25 seconds relate to psychological and textual factors. For prospective subsequent analyses of temporal variables in speech, they advise accepting the value "somewhat over 0.1 seconds". The value of 0.1 seconds is the cut-off point used by Butcher (1981) who observes that breaks under 0.1 seconds are mainly stop closures. On the basis of the analysis of frequency, distribution and the phonetic environment of breaks in the range 0.1 to 0.2 seconds, Dankovičová (1992) suggests a minimum pause duration of 0.13 seconds.

1.3.2 Articulation rate variation across speakers and within the speech of an individual

There is a rather widespread belief that something like 'speaker-specific' articulation rate exists, i.e., that each speaker has his own typical articulation rate, with some tending to speak faster than others. In fact, this has been explicitly

claimed only by a few authors (Goldman-Eisler, 1968; Hála, 1962). In some other works (e.g., Deese, 1984; Sedláková, 1989) articulation rates of individual speakers were claimed to be very similar. Sedláková reports an average articulation rate of 5.42 syll/s for a group of 20 Czech speakers, with a standard deviation of 0.69 syll/s.

The issue of what unit to base the measurement of articulation rate on, discussed in the previous section, is particularly crucial for studies on within-speaker variation in rate. In this respect, the pioneering work of Goldman-Eisler needs to be mentioned first. It was her research on speech rate variation that triggered a discussion in the literature about how much articulation rate varies during speech.

Although Goldman-Eisler (1954, 1956, 1968) concentrated on the duration, frequency, distribution and function of pauses in various contexts and speech situations, she also dealt with speech and articulation rates. Her analysis comprised interviews with eight English speakers. Calculation of the rate variation within the speech of an individual was based on 'utterances', which were defined as speech stretches between successive questions posed by the interviewer. The main results were as follows. The frequency distribution of speech rates showed that there was a high degree of variation; central tendencies, however, were very stable and individual differences proved highly significant. She considered the utterance length to be the major factor in the variation of the speech rate:

> "The range of speech rates is relatively great for short utterances, but narrows as utterance grows longer." (Goldman-Eisler, 1968:22).

This, unfortunately, is a simple statistical fact, as pointed out by Boomer (1970:156):

> "Larger samples from any heterogeneous population are less variable in their central tendency than are small samples, and

> [...] their means are more closely distributed about the population mean."

The average speech rate calculated for all individuals separately varied from 2.3 syll/s to 4.3 syll/s and the average articulation rate varied from 4.4 syll/s to 5.9 syll/s. Goldman-Eisler further calculated, on the basis of mean articulation rates for individuals, the overall coefficients of variation[4] for the percentage of the total time spent pausing (45.9%) and for articulation rate (9.1%). However, this is a rather questionable calculation since it smoothes out variation within individual speakers which was higher. From the figures she presents (1968: 25, Tab. 6), we see that the coefficient of variation for articulation rate ranged between 11.5% and 25.1% when looking at individual subjects while the coefficient of variation for the total time spent pausing ranged between 28.2% and 91.1%. Miller *et al.* (1984:217) remarked that, at least for some speakers, there was reasonable variation in the articulation rate. However, despite her own data, Goldman-Eisler concludes that articulation rate plays no significant part in speech rate variation, and that it is pauses that are the crucial factor. In another study (1961b), she argues that articulation rate is a characteristic discriminating between individuals and 'a personality constant of remarkable invariance' (Goldman-Eisler, 1968:25).

Few authors subscribe now to Goldman-Eisler's belief about invariant articulation rate. There is, however, disagreement as to the amount of variation. It is caused partly by the fact that different researchers use speech sections of different length. Goldman-Eisler's 'utterances' have already been described. In their study of French and English, Grosjean and Deschamps (1975), who measured 30-syllable speech stretches, found coefficients of variation 7.5% and 10% respectively, and regarded such variation as not significant. On the other hand, Malécot *et al.* (1972) used interpause stretches and reported variation of 14.5%. They considered

[4] Coefficient of variation (CV) is a measure that combines information about the mean and the standard deviation (SD) from this mean and it is calculated in the following way: $CV = \frac{SD}{Mean} \times 100$.

rate of interpause stretches to have "a striking degree of conformity" (p.246). They studied samples of conversations from fifty French subjects and found that the average rate was 344 syll/min (5.73 syll/s), with two thirds of all interpause stretches spoken at rates above 300 syll/min (5 syll/s) and below 400 syll/min (6.67 syll/s).

To my knowledge, Miller *et al.* (1984) are the only authors who questioned the role of different methodologies in producing different estimates of rate variation. They emphasized that:

> "The data [...] based on averages across longer stretches of speech do not reveal the extent of local variation in speaking rate within a given utterance of a given speaker." (Miller *et al.*, 1984:218)

They demonstrated this simple statistical principle through the reanalysis of the same (English) material used by Grosjean & Deschamps (1975). Instead of using 30-syllable speech stretches, they chose interpause stretches. The results were quite different from the original. The mean articulation rate across speakers and interpause stretches was 4.97 syll/s, yielding a coefficient of variation of 27%, which is a considerably larger variation, compared to the originally calculated 10% (from the mean of 5.17 syll/s). Miller *et al.* summarized that during a given response, the articulation rate did not gradually increase or decrease, but changed course a number of times. This general pattern was typical of the all 30 speakers in the corpus. They concluded that articulation rate can change substantially not only over the course of an entire interview or even over a single response during an interview, but between adjacent interpause stretches. Their conclusion was that variation of articulation rate is substantial both within and across speakers.

The same generalization was made by Rietveld and Eefting (1988) who examined radio interviews and measured articulation rate across interpause stretches. They found a range of variation between an average syllable duration of 255 ms (3.92 syll/s) and 126 ms (7.94 syll/s).

In summary, there does not seem to be any single solution as to what unit to measure in research on articulation rate variation. This is related to the lack of clarity of the theoretical status of articulation rate. Without a proper theoretical foundation, it is difficult to justify employing any particular unit.

1.3.3 Speech rate, articulation rate and pauses

The issue of how articulation rate and pausing relate one to another was addressed after Goldman-Eisler's research again years later. Butcher (1981) claimed that the way in which pausing and articulation rate may alter as a function of speech rate seems to be more complex than suggested in the past. He takes issue with Grosjean (1980), who argues that speech rate variation is due to varying pause frequency and claims that articulation rate and pause duration are hardly altered. According to Butcher, this is true of slower speech rate. Normal speech rate seems to involve the use of a minimum number of pauses of maximum length which is resistant to reduction. Higher speech rate is achieved by raising articulation rate and reducing the duration of pauses. Butcher's analysis was based on reading in German. With respect to articulation rate, his observations differ from Bártková (1991). She investigated speech rate, articulation rate and pausing in French read texts in order to improve a speech synthesis model for French. She claims that speech rate and articulation rate increased roughly to the same extent from slow to normal, and from normal to fast reading tempo. [5]

[5] Speech rate and articulation rate in Bártková (1991) are more than twice as high as the values presented by other authors for French (see the discussion in 1.3.2). Slow articulation rate was 10.04 syll/s, normal 12.97 syll/s and fast 15.71 syll/s. I suspect this to be due to a much lower minimum pause duration than usual but it is difficult to judge since the information about the minimum cut-off point is missing in the text.

1.3.4 Speech rate and articulation rate variation in relation to independent variables

A number of independent, or extralinguistic variables have been studied in the phonetic literature in relation to articulation rate and/or speech rate, namely, age, task, degree of preparation and, less frequently, sex, occupation, emotional state, and interviewer's behaviour.

Studies which examined more than just one independent variable are rare. One of the exceptions is the analysis of conversation of French adult speakers carried out by Malécot *et al.* (1972). They concentrated on the effect of several independent variables on articulation rate and syllabic length of interpause stretches, and found that only age and occupation were significant for articulation rate. Teachers and executives were among the slowest speakers, speech of students and professionals was faster. Articulation rate was not sex-dependent although women produced interpause stretches more variable in rate than men. No significant effect was found for degree of formality, attitude, subject matter or emotional state.

The study by Malécot *et al.* (1972) is one of relatively few which included age into the investigation of articulation rate as a separate factor (a large number of studies on rate of speaking lack information even about subjects' age). They reported a progressive drop of 0.33 syll/s in articulation rate from young to older adult speakers. For the 'young' speakers (20-29 years), the average rate was 5.95 syll/s, for the 'middle-aged' speakers (30-49 years) 5.65 syll/s and for 'older' subjects (50-69 years), the rate was 5.52 syll/s. Some further evidence that elderly people speak significantly slower than young adults comes from, e.g., Smith *et al.* (1987) and Amerman & Parnell (1992).

With respect to children and teenagers, we seem to have relatively more data available than for adults. The main finding is that articulation rate in childhood and teenage stages increases with age. Starkweather (1980) presents observations

from one of the earliest studies on development of articulation rate, by Dawson (1929), who asked American children of different age to perform a variety of speech tasks as fast as they could. According to Dawson, rate increases far more quickly at an early age (between age 5 and 8) than later on. With regard to sex, the girls talked faster in the earlier years, consistently so up to age 12. At this point the first reversal occurs and is followed by the sexes alternating with respect to which has the faster articulation rate, until age 20 when boys seem to talk much faster. Starkweather claims that after age 20, the rate of utterance gradually declines.

A study by Kowal *et al.* (1975) partially confirmed Dawson's observations that development is discontinuous. They, however, locate the discontinuity at a later age (between age 9 and 11) and suggest that for the next four years after the age of 11, rate continues to increase. Between the ages of 6 and 14 years, articulation rate increases roughly from 2 syll/s to 4 syll/s. Increase in articulation rate is partnered by an increase in disfluencies (pauses, etc.) and in the length of utterance. The disparity between the two studies might be partly due to the fact that Kowal *et al.* (1975) investigated a different type of speech - spontaneous speech.

Fletcher (1972) measured articulation rate in a diadochokinetic task (DDK-task, i.e., repeating sequences of identical syllables as fast as possible) and found that articulation rate increases from about 4syll/s at the age of 6 years to 6 syll/s at the age of 13 years. This is nearly twice the rate in spontaneous speech found by Kowal *et al.* (1975).

Slis *et al.* (1988) also investigated whether articulation rate in spontaneous speech and maximum articulation rate in DDK-tasks are age-dependent and how rate is related to utterance length. Their analysis, based on speech from children aged 5, 7, 9 and 11 years, showed that both utterance length and articulation rate in spontaneous speech are positively correlated with age. Furthermore, age and utterance length were independent in their effects on articulation rate. The

difference between articulation rate in spontaneous speech and maximum rate in the DDK-task was not very large, the former being slightly lower.

The question why children and elderly people speak more slowly than young adults is open to speculation. As to children, the general view is that the reasons are the limitations posed by neuromotor or linguistic immaturity and a lesser degree of communicative experience (Amerman & Parnell, 1992). Whether the slower rate in older adults is due to decline in neuromotor capabilities is debatable. Birren *et al.* (1979, 1980) pointed out that slowing of performance is one of the most characteristic changes in old age, affecting all stages from motor output to sensory input. However, some alternative explanations have been put forward, in particular, decline in cognitive efficiency and 'speed/accuracy tradeoff'. Welford (1977) and Kent and Burkard (1981), among others, believe that a tendency towards increased caution, concern with accuracy, or more careful monitoring of motor output may be responsible for preferring slower execution of speech. All these explanations seem plausible but the exact contribution of each of these remains a mystery.

Articulation rate is obviously limited by physiological systems. Such factors as the intrinsic speed of articulatory activity and the transmission time for motor impulses to the speech muscles are of primary importance in this respect (Dalton & Hardcastle, 1977). It has been suggested that we talk close to the limits of our ability:

> "The evidence for this belief comes from the relative lack of individual or situational variation in rate and from the fact that the fastest reaction times and diadochokinetic rates are closely parallel to average speech rate[6]" (Starkweather, 1980:161).

This claim, not supported by any data from the author himself or by references to any other studies, is questionable, especially in its first part. The phenomenon of

[6] This corresponds to articulation rate in standard terminology.

individual variation in rate is problematic, as will be seen below, and the concept of 'situational variation' is not defined. If variation due to different tasks is meant to be included under this term, there is extensive evidence for considerable situational variation.

Task and degree of preparation are probably most commonly investigated independent variables. Goldman-Eisler (1961b) did not find any difference in articulation rate with respect to distinct cognitive operations (the description of cartoon stories and the interpretation of their meaning) but found an effect on rate of the amount of practice (repetition of speech). However, her methodology involves dubious measurement techniques, as noted earlier, and her results may be affected by this.

The degree of preparation was also found to be important by Barik (1977) in his analysis of temporal characteristics across different types of speech material (prepared speech, semi-prepared and spontaneous speech) in English and French. The same tendency was found in both languages: more prepared speech had a higher speech rate. In a prepared lecture, the speech rate was 3.16 syll/s for English and 3.48 syll/s for French. The speech rate in a spontaneous story by English speakers was 2.5 syll/s and by French speakers, 2.19 syll/s. For articulation rate, however, the tendency was reversed, prepared speech being slower and spontaneous speech faster (4.28 syll/s for English and 4.27 syll/s for French lecture; 4.87 syll/s for English and 5.05 syll/s for French storytelling). The inverse relationship between speech rate and articulation rate was explained rather obscurely in terms of 'extraneous factors' but it is clearly due to a different amount of pausing in these two tasks. In the lecture in English, pauses took up 25% of the total speech time, in French 18%. In an English story it was 49% and 56% in a French story.[7]

[7] The minimum duration which Barik recognizes as a pause is 0.6 s which is a higher value than in most studies of this kind (for a survey see Kowal *et al.*, 1983). Leaving out pauses of shorter duration must have affected the results.

Kowal *et al.* (1983) compared the use of time in storytelling and interviews as reported for different languages in the literature. The comparison, based on weighted means to balance out differences in methodologies, showed no significant difference in articulation rate (5.17 syll/s and 5.26 syll/s respectively), but considerable difference in pausing (pauses took up 33% of the total time in storytelling and 17% in interviews).

Fónagy and Magdics (1960) investigated material that was more diverse in nature, sports commentaries and recitations of poems in Hungarian. It is not very surprising that sports transmissions were considerably faster than poetry recitations (13.8 sounds/s vs. 9.4 sounds/s).

Sports commentaries in Czech are faster than spontaneous speech, according to the following two studies (data are available only for speech rate). Bartošek (1974) found 5.83 syll/s to be the average and 6.67 syll/s the maximum speech rate in sports transmissions. Romportl (1958) reported a range between 3.58 syll/s and 4.17 syll/s for normal speech rate in his examples of spontaneous speech, with the maximum of 4.83 syll/s.

Differences between articulation rate in reading and spontaneous speech (reading being faster) were observed, for instance, in Lehtonen (1979) and Butcher (1981), the latter reporting the value of 6.13 syll/s in reading and 5.26 syll/s in spontaneous speech but no significant differences between two types of spontaneous speech (reminiscence and retelling a story). A faster articulation rate in reading than in spontaneous speech was also found by Strangert (1993) for Swedish (5.8 syll/s as opposed to 4.2 syll/s).

The effect of speaker's attitude and emotional state on articulation rate has not been investigated systematically to such an extent that a clear picture, supported by quantitative research, would emerge. Some studies, based on an impressionistic analysis (e.g., Petřík, 1938; Crystal, 1969; Cruttenden, 1986), suggested that

emotional state does affect speaking rate. However, a study by Malécot *et al.* (1972), mentioned earlier, did not find evidence for such an effect. This study is one of few experimental works dealing with this topic. More research is clearly needed and there are indications of some work in progress (Arnfield *et al.*, 1995).

Lack of a sufficient body of research applies also to the effect of sex on articulation rate. The results available so far are not consistent. For instance, Malécot *et al.* (1972) claims absence of such an effect in his study of French; Deese (1984), on the other hand, found for English that women speak on average somewhat faster (5.82 syll/s) than men (5.48 syll/s). Whether the difference in the results of these two studies are due to the different languages involved is an open question.

Among other factors believed to affect speech rate was the interviewer's behaviour, namely his own rate and the frequency with which he interrupted the interviewee's speech (Webb, 1969). Another extrinsic factor, suggested by Black (1950), is the size and reverberation time of the room but not its shape. Black found that rate of reading aloud was significantly slower in a larger and less reverberant room.

1.3.5 Other determining factors

Some researchers have questioned whether articulation rate is dependent on the length of the speech stretch for which it is measured. It has been repeatedly shown that syllables are produced consistently faster in longer utterances[8] than in shorter ones (e.g., Bártková, 1991). Fónagy and Magdics (1960) demonstrated this tendency for the 'phrase' in Hungarian, defined as a rhythmical unit, as a sound

[8] The term utterance is used in a very general sense here and it stands for speech sections of various kind (interpause stretches, clauses, etc.).

sequence governed by a stressed syllable.[9] They express the dependence of articulation rate on the length of phrases by means of exponential functions and argue that the tendency to pronounce short and long phrases in an approximately equal time is related to a more general rule for speech production, a 'law of equalisation'. Malécot *et al.* (1972) worked with interpause stretches and noticed that in their material, short interpause stretches deviate from the mean articulation rate twice as much as the long stretches. Starkweather (1980) suggests two alternative explanations for this observation: (i) a difference in the amount of information load (short utterances contain more information per syllable) and (ii) running out of breath in the case of longer utterances. These explanations are rather speculative and in the case of the latter, the author himself admits that it does not have great plausibility. The explanation is probably more straightforward: shorter stretches of speech are naturally less likely to be close to the overall population means than longer stretches.

Romportl (1958) pointed out a variation of articulation rate in relation to the type of a phrase. As an example, he cited parenthetical remarks, which often involve considerably faster or slower rate then the surrounding phrases. Rate changes indicating contrast between grammatical structures were also noted by Crystal (1969).

In his study of sentences, phrase rate and pause characteristics of oral reading in English, Brubaker (1972) found that sentence rate increased progressively towards the end of the text and the second phrase in the sentence was consistently faster than the first. As a possible, but rather speculative, explanation, he suggests that rate increases (and pauses decrease) as uncertainty decreases. As one proceeds through the text, the meaning becomes more obvious and predictable. The data are, however, not easily comparable with other studies on sentence/phrase rate because of the measure used - the number of words/min - and there is always a

[9] We would call it probably the 'foot' or the 'phonological word' nowadays.

question of whether the same results would have been obtained if a different measure had been used.

Recent studies from the USA, Sweden, Holland and Japan dealing with temporal organization in speech in connection with speech synthesis provide some new information about articulation rate and its variation within a speaker. Crystal & House (1990) claimed that in their speakers' reading of an English passage, the variation of articulation rate, measured across interpause stretches, was a function of the phonological characteristics of the interpause stretch, specifically, the average number of phones per syllable and the proportion of either stressed phones or stressed syllables in the interpause stretch.

In a similar study of Dutch, Eefting (1990) used stretches produced in their original context as opposed to stretches embedded into a carrier phrase, and found a significant correlation between articulation rate and the phonological characteristics of the stretch only when the original context was absent. Eefting suggested that variation in articulation rate is not determined solely by phonological factors but reflects also the intention of the speaker to accommodate to the audience, i.e., to facilitate the communication process.

Crystal and House's (1990) suggestion that rate variation depends solely on phonological factors was directly challenged by Nooteboom and Eefting (1994). In a methodologically similar study of (presumably Dutch)[10] newsreading by a professional speaker, they showed that rate variation of the same phrases was considerably different when placed in the original, proper textual context, as opposed to a neutral context. Moreover, the authors demonstrated that deviations from the intended rate were noticeable by listeners and negatively affected perceived naturalness. They also noticed that listeners were sensitive to variations within the phrase and consequently they suggested that the unit of articulation rate is not the interpause stretch but rather a unit below this level. As one of the

[10] Language was not specified in the paper.

possible contextual determinants of rate variation within the stretch they mentioned 'newness' and 'givenness' of information.

Fant *et al.* (1991a) also concluded that the factors underlying variations in local articulation rate include text-specific and speaker-specific factors. They investigated articulation rate variation in the reading of several Swedish sentences, measuring across phrases demarcated by pauses and 'other syntactic boundaries'. The authors reported both phonological and syntactic influences of text, specifically, the relative density of stressed syllables, the occurrence of major clause boundaries and word class (content vs. function words). With respect to speaker-controlled effects, articulation rate reflected whether or not a phrase attained focal attention. In general:

> "A common pattern within a sentence is that the mean phoneme duration starts low, rises to a peak and decays. In other words, a deceleration followed by an acceleration of local speech rate." (Fant *et al.*,1991a:256)

The study of the literature on articulation rate variation reveals that the term 'local articulation rate' is not clearly defined. Den Os (1988), for instance, characterized it as changes in rate resulting from, for example, factors determining final lengthening. She contrasted 'local rate' with 'global rate', which she describes, again rather vaguely, as changes in overall speaking rate. In his discussion of duration models, Carlson (1991) characterized 'local articulation rate' as a rate change inside the phrase or the sentence. He claimed that in most approaches to duration models, it is assumed that the speech tempo[11] is constant during a sentence or clause (no references were given), and argued, on the basis of studies involving duration models, that this is not true. He says:

[11] Carlson did not distinguish between speech rate and articulation rate in his paper and his 'speech tempo' would be 'articulation rate' in standard terminology.

> "It is interesting to note that in all these studies, the phoneme durations have a standard deviation of around 40 ms. After some kind of model is applied we typically get an error with a standard deviation of 25 ms." Carlson (1991:244)

Carlson (1991) noted further that when comparing a prediction of duration to natural speech, the prediction error is often found to be a function of time and suggests that this result may be a manifestation of local articulation rate.

Fant *et al.* (1991a) is one of few studies concerned with local rate variation. It is unfortunate that the unit across which their measurements were carried out seems arbitrary and is not well defined. However, there is no generally accepted definition of local articulation rate with respect to what unit it is bound to.

A possible way to address this problem is to relate articulation rate variation to word duration. Here the dichotomy between content and function words is again claimed to be significant (e.g., Kaiki *et al.*, 1990; O'Shaughnessy, 1995). However, Carlson (1991) suggested that the syntactic position of the word, rather than word class, might be the crucial factor. O'Shaughnessy (1995) provided evidence for both the effect of word class and syntactic position. He investigated factors determining word duration in samples of spontaneous English speech. In fluent speech, he found consistent durations, especially for monosyllabic function words. Stressed content words varied more in duration, especially in prepausal position. He also observed that words at the beginning of a syntactic unit tend to be shorter than words at the end of the unit (even discounting prepausal lengthening). Phonological properties also appeared to be relevant at the word level. O'Shaughnessy argued that a large amount of the variability in word durations could be explained in terms of the number of phonemes in the word. Crystal and House (1990) reported that the duration of stress groups had a quasilinear dependency on the number of syllables and the number of phones.

Considering the discourse level of speech, it has previously been suggested that speakers may signal new or important information in their discourse by a decrease in articulation rate whereas predictable or given information will be associated with a faster rate (e.g., Hunnicut, 1985; Fowler & Housum, 1987; O'Shaughnessy, 1995). However, Eefting (1991) and Eefting and Nooteboom (1991) pointed out that information value seems to be closely related to accentuation and that in none of these studies was information value varied independently of accentuation. It is thus not clear to what extent each of these factors has a separate effect on word duration. To examine this issue, Eefting conducted an experiment on content words from read Dutch and found that word duration depended on accentuation (accented words were longer) but not on information value. Accentuation also seems to be more important to intelligibility of speech than whether the information had been used before in conversation (Hawkins & Warren, 1994).

Associative context (Shields & Balota, 1991) and repetition (e.g., Fowler, 1988) have also been shown to affect word duration. Shields and Balota demonstrated that in English mean word durations were shorter in the repetition condition, compared to the associatively related condition (in which the target word was semantically related to some other earlier word in the sentence), and this in turn produced shorter durations compared to the unrelated condition. Fowler (1988) examined shortening of word durations in a variety of communicative contexts and found a shortening effect in prose reading as opposed to reading a word list, in which no shortening of repeated lexical items occurred. Moreover, she discovered that shortening was more pronounced in spontaneous speech with a silent listener present than when the same speaker read back a transcript of the speech. McAllister *et al.* (1994) suggested that this result might be due to the presence of a listener and considered it as supporting a collaborative view of the conversation, which assumes that speakers attune their productions to the state of knowledge of their addressee. They reported further evidence for this view on the basis of their comparison of monologues and dialogues. They observed that words were generally shorter in dialogues than in monologues, and the reduction in repeated word durations was more pronounced and immediate in dialogues, as

opposed to monologues, in which decrease in duration across successive repetitions was more gradual.

Campbell (1991) provided evidence, on the basis of a recorded narrative in English, that not only repeated nouns shorten but also their subsequent related pronouns. He claimed that:

> "Both antecedent nominal units and subsequent pronominal references to them [anaphoras] can be seen to reduce as the utterance progresses, but with resetting occurring after an interval of time and at major breaks in the narrative of the text." (Campbell, 1991: 286)

A different kind of contextual effect has been investigated by Weismer and Ingrisano (1979; reported in Cooper & Paccia-Cooper, 1980), namely, the effect of emphatic stress on the duration of other words within the sentence. They claim that emphatic stress not only increases the duration of the emphasized word but also, when placed near the beginning of the sentence, produces shortening of word duration toward the end of the sentence.

Any discussion of determinants of word duration would be incomplete if phrase-final lengthening were not mentioned. I have already touched upon this topic indirectly when I remarked on research that excluded words in final phrase/sentence positions from their investigation to prevent a confounding effect of this phenomenon (e.g., O'Shaughnessy, 1995). The amount of research conducted on phrase-final lengthening is considerable. The most common perspectives that this research has taken include: relation of phrase-final lengthening to syntactic structure (e.g., Klatt, 1975; Cooper *et al.*, 1978; Cooper & Paccia-Cooper, 1980; Selkirk, 1984), to prosodic structure (e.g., Beckman & Edwards, 1990; Gussenhoven & Rietveld, 1992; Ferreira, 1993; Hofhuis *et al.*, 1995), and the location of the lengthening effect within the final word (e.g., Cooper & Paccia-Cooper, 1980; Edwards & Beckman, 1988; Fant *et al.* 1991b;

Berkovits, 1993, 1994; Lyberg & Ekholm, 1994; O'Shaughnessy, 1995; Hofhuis *et al.*, 1995; Horne *et al.* 1995). Phrase-final lengthening has been attested across a number of different languages (for a survey, see Vaissière, 1983). Although we cannot yet say whether it is really a universal phenomenon (it does not seem to occur, for instance, in standard Japanese, Hallé *et al.*, 1991), the evidence for it is strong and thus phrase-final lengthening needs to be taken into account in studies on word duration and articulation rate variation.

1.3.6 Do languages differ with respect to articulation rate?

A number of studies have attempted to answer the question whether some languages are spoken faster than others. However, in some instances the comparison was based on speech rate, which is, as mentioned earlier, a complex measure of two, by nature very different, phenomena - pausing and articulation rate. Speech rates appear to show a greater cross-linguistic variation than articulation rate (Butcher, 1981). This, in fact, means greater variability in pause duration and/or frequency (see the discussion in section 1.3.3), and relatively stable articulation rate. Since we cannot infer how articulation rate behaves on the basis of speech rate data, the focus of cross-linguistic studies should be on articulation rate.

Lenneberg (1967) claimed that articulation rate reflects a basic rhythm which characterizes the temporal patterning underlying neuro-muscular automatisms. These are based on a physiological rhythm with a periodicity of 6 Hz. He added that:

> "If this rhythm is due to physiological factors rather than cultural ones, it should be then present in all languages in the world." (Lenneberg, 1967:118)

Although it is clear from the earlier discussion that articulation rate is affected by a number of linguistic and extra-linguistic factors and they interact in a complex way, some universal underlying mechanism, physiological or cognitive, is likely

to exist. How to extract it from the 'surface' form, the production of connected speech, seems presently to be a task beyond our power.

Study of the literature on articulation rate does not give us a clear answer to the question whether languages differ in articulation rate. But the prevailing impression is that of considerable similarity. The most valuable source of information in this respect are cross-linguistic studies since they use the same methodology for the languages they compare. As I shall illustrate later, this point is very important. Unfortunately, cross-linguistic works are rather sparse. Data from some of them have already been mentioned in my discussion of articulation rate in relation to independent variables (section 1.3.4). For instance, the experiment of Barik (1977), compares English and French and reports a very similar articulation rate for the two languages, within the same type of material. A somewhat larger discrepancy was, however, found by Grosjean and Deschamps (1972) for interviews. Articulation rate was 5.29 syll/s in French and 4.83 syll/s for English.

Osser and Peng (1964) did not find differences in rate between American English and Japanese. Den Os (1988) investigated Dutch and Italian (reading). He asked his speakers to read several texts at normal and fast tempi and found no significant difference in articulation rate between the two languages in both tempi (normal tempo: 6.1 syll/s for Dutch and 6.4 syll/s for Italian; fast tempo: 8.1 syll/s and 8.4 syll/s respectively).

Information about articulation rate in different languages can be obtained also from analyses based on only one language. Some data for German, Czech and Hungarian have already been mentioned in section 1.3.4 and I shall provide some additional data here. For German, Henze (1951) reported the overall mean articulation rate of 5.55 syll/s and Butcher (1981) 5.7 syll/s. For French, we have more data from Duez (1982), who investigated political interviews (5.7 syll/s) and casual interviews (5.2 syll/s). The data correspond relatively closely to the data

from interviews in French discussed in other papers; Malécot *et al.* (1972) reported 5.73 syll/s and Grosjean and Deschamps (1975) 5.29 syll/s. Further indirect cross-linguistic comparison was provided by Kowal *et al.* (1983), who analysed previously published data for five languages (English, Finnish, French, German and Spanish). The rate for these five languages varied between 4.09 syll/s (Raupach, 1980; German) and 6.08 syll/s (de Johnson *et al.*, 1979; Spanish). These are the extreme figures; the majority of the data lies between 4.5 syll/s and 5.5 syll/s. Kowal *et al.* concluded that differences in articulation rate between languages are negligible compared to differences due to speech task. (cf. Barik, 1977).

But could we really claim on the basis of the data available that articulation rate in languages is similar? There are two problems with this. Firstly, a comparison of articulation rates for different languages investigated with different methodologies can give us only an approximate picture, certainly not very reliable. Secondly, it is not clear where the boundary between 'different' and 'similar' lies; for instance, does the articulation rates of 4 syll/s and 6 syll/s count as considerably different, 5 syll/s and 6 syll/s as just different, and 5.5 syll/s and 6 syll/s as similar?

The issue of comparability of data collected with different experimental settings involves primarily the following points: differences in (i) minimum pause duration (cut-off point), (ii) the measure employed, and (iii) the length of the speech section on which the measurements are based. Points (i) and (iii) were already discussed in section 1.3.1; I shall concentrate here on various ways of expressing articulation rate.

Measures of rate employed in the literature include number of words per minute, number of syllables per minute or per second and number of sounds per second. The measure of words per minute was used in some earlier studies but nowadays it seems to be generally abandoned. There is a good reason for this, as demonstrated, for instance, by Lehtonen (1979), who found that Finnish was faster

than English if syllables per second were used as a measure, but that it was slower if words were used as counting units. The differences seem to be, at least to some extent, explained by the tendency of Finnish words to contain more syllables than English words. A similar principle, however, applies to the measure of syllables as opposed to sounds/phonemes per unit of time. A language with a simpler syllable structure is likely to have a faster articulation rate than a language with a more complex syllable structure. This has been demonstrated by den Os (1988). He found no significant difference between Dutch and Italian when articulation rate was measured in syllables per second but when sounds per second were employed, Italian was significantly slower. When trying to compare data on articulation rate from different studies, it is possible to recast one measure into another[12] on the basis of statistics concerning the average word duration in the language in question (in syllables or in phonemes) and information about the most frequent types of syllables. However, such information may not always be available and, even when it is, the picture we obtain after recalculation is still only approximate.

Nowadays, the majority of authors use the measure of syllables per second. Differences in the complexity of syllable structure are less drastic among languages than differences in word length. Using syllables rather than segments, despite some variability in their complexity, can be justified theoretically. I shall return to this particular issue in Chapter 2.

Finally, I would like to come back to the question of defining 'similar' and 'different' for articulation rate. No objective criterion has been suggested in the literature. We could certainly test differences in articulation rate perceptually across different languages. However, when perceiving continuous speech, we seem to be influenced by the amount of pausing in speech (Ofuka, 1996; mentioned in Roach, in press), which would suggest that in perception of continuous speech, it is speech rate rather than articulation rate that is crucial.

[12] As was done by Kowal *et al.* (1983) in their analysis.

Another difficulty with perceptual testing is that it is still unclear whether we perceive rate in foreign languages similarly as in our own native language. Our perception of articulation rate might be influenced also by other factors than articulation rate itself, for instance, rhythm, which is closely related to articulation rate, or differences in linguistic structure (in morphology, in word order, etc.). However, there is some evidence that this might not be the case (Vaane, 1982). In any case, before we can make stronger claims about similarities or differences between languages with respect to articulation rate, more data are needed. Large databases, such as EUROM-1 (Chan, 1995) and BABEL (Roach *et al.*, 1996) might be of great help here.

1.4 AIMS AND OUTLINE OF THE THESIS

As I indicated at the beginning of this chapter, articulation rate is a temporal phenomenon rather neglected in phonetic research. The survey of the literature showed that only some of the issues listed in section 1.2 have been addressed. So far neither differences in articulation rate between speakers nor rate variation within the speech of an individual have been investigated in sufficient detail for any language. The number of studies which go beyond providing a simple measure of the overall articulation rate for the speech of an individual is very small. Moreover, comparison of results from these studies is often impossible since the methodologies used are different. This reflects the fact that it has thus far been unclear what the appropriate method is to examine variation in articulation rate. I believe that this problem is largely due to our uncertainty as how to incorporate this phenomenon into linguistic theory.

In my thesis, I attempt to contribute to our understanding of principles on the basis of which articulation rate functions. The area which seems to me central in this respect, and which became the main topic of this thesis, is the nature of articulation rate variation. In the absence of a theoretical explanation for such

variation, the main question is whether variation in articulation rate has a specific domain, and, if so, (i) how such a domain can be described linguistically, (ii) to what extent articulation rate varies, (iii) what the internal structure of rate variation is within that domain, and (iv) what the main factors are which determine articulation rate variation. The thesis addresses all these questions. It presents an experimental phonetic investigation of variation in articulation rate from the perspective of speech production. I adopted a naturalistic approach to gathering data as opposed to a controlled-experiment setup (Garman, 1990), i.e., the data comes from relatively natural, contextually embedded speech. This seems to me necessary at this stage since, firstly, our up-to-date knowledge about articulation rate variation is too limited to be able to form strong enough hypotheses needed for controlled experiments, and, secondly, I believe that it is very valuable to have data from naturally occurring continuous speech.

The language this thesis is concerned with is Czech. This language is closest to me as it is my native language. But there are other motivations behind investigating articulation rate variation in Czech. For instance, it makes it possible to make claims about rhythm in Czech, in particular, whether Czech has a stronger tendency towards stress-timing or syllable-timing. In a strongly stress-timed language, the articulation rate of the word would be expected to vary linearly with the word's size. In a strongly syllable-timed language, the rate would stay more or less constant, irrespective of the word's size. Traditionally, Czech is claimed to be a syllable-timed language (Palková, 1994); however, some authors prefer to characterize it as having both syllable- and stress- timing (e.g., Daneš, 1957; Bolinger, 1965). It is also of interest to see whether the relationship between the word's size and its articulation rate is affected by vowel quantity, a contrastive feature of Czech.

Apart from pursuing the main aim of this thesis, i.e., to capture the nature of variation in articulation rate in Czech, I hope to encourage the beginning of a theory of articulation rate variation, including its relation to speech planning and

speech production. This thesis also aims to provide a model against which further studies, including cross-linguistic, could be carried out. Providing the method devised here for investigating articulation rate variation with a theoretical justification, this work strives to avoid arbitrariness in methodology, so typical in the research on articulation rate reported in the literature. In this way, I believe it can open space for further, extensive research on articulation rate variation across languages. A potential for speech technology research has been already mentioned at the beginning of this chapter. Last, but not least, the thesis offers a pool of data on a language in which articulation rate has not been analysed to any comparable extent.

The thesis reports on two separate studies. Study 1 is concerned with the domain of articulation rate variation. Specifically, it examines three candidate domains, the interpause stretch, the clause and the intonation phrase. Since the intonation phrase turned out to be the successful candidate, articulation rate variation within this domain was investigated in more detail. This part can be considered a pilot study on articulation rate variation within the intonation phrase, including the issue of determinants of articulation rate variation. Study 1 is covered by two chapters; in Chapter 2, experimental design is discussed, and in Chapter 3, the results of the experiment are reported. The work presented as Study 1 is based on two of my previous works: a chapter from my M. Phil. thesis (Oxford, 1994) and the paper "The domain of articulation rate variation in Czech", published in the Journal of Phonetics in 1997. The paper represents an extended and revised version of my M. Phil. work.

Study 2 builds on the results of Study 1. It is solely concerned with the patterning of articulation rate variation within the intonation phrase. The main aim was to test the results from Study 1 on a larger database, both in the sense of the number of subjects and the number of different tasks. However, an important part of Study 2 is also an examination of a number of linguistic factors presumed to have a significant effect on articulation rate. Chapter 4 presents notes on methodology.

Chapter 5 provides a survey of elementary results, including the degree of cross-subject consistency in articulation rate variation when reading the same text. Chapter 6 contains the results of the statistical analysis of articulation rate patterning and its determinants and is the core of the thesis.

Chapter 7 contains an attempt to design a mathematical model that would best fit the actual values of articulation rate observed in Study 2, given the explanatory variables investigated in Chapter 6. It represents a step towards a quantitative representation of articulation rate variation, on the basis of which further work could be carried out to allow predicting articulation rate.

Finally, Chapter 8 summarizes and discusses the findings from the two studies. It suggests that articulation rate does have, at least in Czech, a specific domain of variation, and that this domain is the intonation phrase, within which articulation rate reflects a pattern of progressive slowing down. The results are discussed in relation to previous research on articulation rate and in relation to theories of speech planning and theories of speech rhythm.

CHAPTER 2

STUDY 1: EXPERIMENTAL DESIGN

2.1 INTRODUCTION

In this chapter I shall explain the main hypotheses and methodology of Study 1, which was concerned with the question whether variation in articulation rate has a particular domain.

As I explained in section 1.2, by domain of articulation rate variation I mean a unit within which articulation rate is organized. I also outlined two different ways of organization - constant rate within a unit as opposed to patterning of rate. Methodologically, the issue of domain of articulation rate variation can be approached either by (i) measuring articulation rate first and only subsequently by looking for changes in rate variation, from which a domain would be inferred, or by (ii) hypothesizing first what units of speech could be possible candidates for domain of articulation rate variation, and then testing the hypotheses by measurements. In my study, I adopted the latter approach.

One hypothesis could be that some characteristic change in articulation rate happens, say, every twenty syllables. However, language is not a simple string of syllables but a complex structure. It is thus highly unlikely that a domain of articulation rate variation would be stretches of a regular number of syllables, without any relation to the linguistic structure. It is much more likely that the domain, if it exists at all, coincides with some linguistic unit. One should be aware that with our particular choice we already make assumptions about the underlying mechanisms controlling variation in articulation rate. So, for instance,

if we choose the clause as a possible candidate for the domain, we assume that syntactic processing may control also processing of articulation rate.

A decision has also to be made about how to measure articulation rate, i.e., what should be the unit across which articulation rate variation should be measured; should it be measured for each syllable, word, sentence? The choice will clearly depend in part on the size of the candidate domain within which we want to examine articulation rate; our unit should obviously be smaller than this domain. But our decision will also be driven by assumptions about speech processing in a similar way as in choosing candidate domains.

In this study, the candidate domains were the interpause stretch, the clause, and the intonation phrase. The unit used for measurements of articulation rate was the phonological word (hence mostly 'word'). The justification for this choice and the choice of the candidate domains is discussed in this chapter.

In my study on domain of articulation rate variation I started with the null hypothesis[1]:

Hypothesis 1.1 (null hypothesis):
Articulation rate is constant within the candidate domain.

This means that any changes in rate occur only across, never within, units. Given the results of past research on word duration, discussed in Chapter 1, the predicted outcome is that this hypothesis is false. An additional reason for this prediction is the likelihood of phrase-final lengthening occurring in some, if not all three candidate domains. As I mentioned in section 1.3.5, evidence for phrase-final lengthening exists for both the clause and the intonation phrase. The term pre-pausal lengthening, which is also common in the literature, suggests that it may

[1] The number before the dot represents the number of a hypothesis; the number after the dot indicates whether the hypothesis relates to Study 1 or Study 2.

be, in fact, the interpause stretch that determines final lengthening. Testing the null hypothesis might shed some light on the still controversial question of which unit of speech governs phrase-final lengthening.

Given a rejection of the null hypothesis, my subsequent hypothesis was:

> **Hypothesis 2.1:**
> **Articulation rate variation within a candidate domain reflects a certain systematic pattern, which is the same in all speakers.**

This means testing whether the position of the phonological word within a candidate domain significantly affects the word's articulation rate. It needs to be pointed out that there does not have to be only one domain of articulation rate variation; each candidate domain may have its own characteristic systematic pattern.

Although in this study the primary concern was the structure of articulation rate variation, i.e., the importance of the factor of the word's position in a candidate domain, the factor of size of the phonological word was also considered as a possible determinant of articulation rate. As explained in section 1.4, considering this factor allows us to test the findings reported for other languages on Czech. The hypothesis is as follows:

> **Hypothesis 3.1:**
> **The size of the phonological word significantly affects the word's articulation rate in all speakers.**

2.2 MATERIAL

The material used in this study consisted of three samples of elicited spontaneous Czech speech from three speakers. The samples were chosen from a much larger corpus which was assembled in 1988 by the Phonetics Institute at Charles University, Prague. Speakers participating in the recording of this corpus were told that the Institute was collecting a database for the study of the standard of contemporary Czech as spoken by the younger generation.

The task used in the present study was to talk for two minutes about practically any topic. A picture taken from a children's book was supplied as a catalyst, in case the subject had no idea what to talk about. Speakers were not informed about the nature of the task until the very beginning of the recording itself; in this way any advance preparation was avoided. The three subjects used in this study chose to produce a narrative, which was inspired by the picture. Their speech was more fluent than the speech of the majority of other subjects recorded, who simply described the picture. A comparable task was an important consideration given that a number of authors reported an effect of speech task on articulation rate (section 1.3.4). Reasonable fluency and the impression of naturalness were the other criteria for the choice of subjects for the present study. The study of articulation rate variation would be difficult to carry out on speech in which every second word is followed by a pause.

2.3 SUBJECTS

Apart from the above criteria, the selection of subjects was guided by the effort of having as homogeneous a group as possible with regard to age, sex, education, familiarity with the interlocutor, and dialect. There are several reasons for controlling for these extralinguistic factors. Firstly, some of these factors, particularly age, sex, and the degree of familiarity with the interlocutor, were

reported in the literature as significant factors affecting the temporal structure of speech (see section 1.3.4). Secondly, factors, such as education and dialect, might be relevant in this respect too, but the literature does not offer any answer as to how and to what degree. It was thus safer to keep these factors the same for all subjects as well.

The three speakers used in my experiment were all female, native speakers of Czech, aged between 20-25 years. They were students at Charles University, Prague, and had no strong regional dialect.[2] Neither of them knew the recording person before.

2.4 RECORDING PROCEDURE

The recording took place in the recording studio of the Phonetics Institute, Charles University, Prague with a reel-to-reel tape recorder TESLA B113 and a microphone SHURE 5855B. During the recording, only one additional person was present, to give instructions; this was a member of the Institute.

[2] Dialectal differences between various regions in the Czech Republic are far from being as strong as in Britain. There are only very few areas (mostly in Moravia) where we could still find a dialect with a specific set of linguistic features different from standard or spoken Czech.

2.5 METHODOLOGY FOR THE ANALYSIS OF ARTICULATION RATE VARIATION

2.5.1 Candidate domains: definitions

As mentioned earlier, three candidates for the domain of articulation rate variation were chosen: the interpause stretch, the clause and the intonation phrase. These units are defined in the following way.

2.5.1.1 Interpause stretch

Interpause stretch is a stretch of speech bounded by consecutive pauses, where a pause was any gap in speech 130 ms or longer which was not a stop closure. This value was taken as a reasonable minimum pause duration on the basis of an experiment in which a similar type of material was used (Dankovičová, 1992). As discussed in section 1.3.1, a value between 100 ms and 150 ms also seems to be supported by other research on minimum pause duration (e.g., Heike *et al.*, 1983). Silent and filled pauses (i.e., pauses filled with 'er', 'ehm' or similar non-linguistic phonetic material) were treated as a single group.

The word 'run' has been used for this unit by some authors (Miller *et al.*, 1984); some others employed the term 'utterance' (e.g., Malécot *et al.*, 1972; Butcher, 1981) or 'speech chain' (Bártková, 1991). Since the word 'utterance' is used in the literature in a number of different senses, 'interpause stretch' is a clearer, self-explanatory term.[3]

[3] Apart from being defined as the interpause stretch, 'utterance' has also been used to refer to the period of speech of a subject between two successive contributions of an interviewer (Goldman-Eisler, 1968). In the Czech phonetic literature (e.g., Daneš, 1957), the utterance is defined as a "smallest whole of speech, the fundamental unit of communication [...] it is delimited by intonation and its own organization as a part of a given context (this organization is based on the differentiation between 'the theme' and 'the rheme')" (Daneš, 1957:139).

2.5.1.2 *Clause*

The clause was defined according to traditional analyses of Czech syntax (Petr, 1986), i.e., a part of a text based around a predication, where the predicate is a finite verb. It is thus a unit of syntactic organization higher in hierarchy than the syntactic phrase and lower than the sentence.

2.5.1.3 *Intonation phrase*

The term 'intonation phrase'[4] is used here in a sense traditional in Czech phonetic literature (e.g., Daneš, 1957; Palková, 1994), and it corresponds to the tone unit in British intonational analyses (e.g., Crystal, 1969; O'Connor & Arnold, 1973) and the intonation phrase in the Beckman and Pierrehumbert's work (1986).

The intonation phrase is a stretch of speech determined by the intonation contour in which the most important characteristics are the presence of the nucleus (the pitch prominence or the intonational centre) and the tail (syllables following the nucleus which demonstrate a certain characteristic direction of pitch movement). These two characteristics together form a particular intonation pattern that is labelled in the Czech literature as 'cadence'. However, it has been argued that according to perceptual tests,

> "the phonic centre [the nucleus] cannot really be considered as a sufficiently conspicuous (or unequivocal) language phenomenon to be reliably reflected in the listener's awareness" (Palková, 1974:113).

Thus the definition of the intonation phrase for Czech that Palková offers is the following:

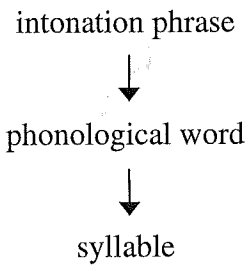
---

[4] In Czech terminology, 'promluvový úsek'.

> "The discourse segment[5] is a group of word-stress groups[6] (in extreme cases, a single word-stress group) which the language-user perceives as a self-coherent intonational whole, most often on the basis of the phonic features of its [final] boundary." (Palková, 1994:162-163; my translation)

The most important boundary markers are generally considered to be the cadence and the pause. The pause on its own, however, is not a sufficient element for boundary determination because its occurrence is not limited to the boundary only; it can occur within a intonation phrase as well (this concerns particularly some hesitation pauses). The cadence is, however, both necessary and sufficient. In addition, a frequent feature of the intonation phrase boundary is claimed to be both in Czech and English final syllable lengthening (e.g., Palková, 1974; Cruttenden, 1986).

According to Czech phonetic literature, the intonation phrase in Czech sits at the top of the prosodic hierarchy:



The intonation phrase consists of one or more phonological words, which are considered rhythmical units in Czech; these then contain one or more syllables.

The stressed syllable of the last phonological word in the intonation phrase also normally carries the nuclear accent, even in emotionally marked utterances and utterances with a marked information structure. In Czech, the word order may be

---

[5] This corresponds to the term 'intonation phrase' in my thesis.

[6] Word-stress group is to be seen as the same entity as the phonological word as used in the present study (for definition see below).

changed in order to ensure that the emphasised word is last in the phrase. (But the shift of the nucleus from its usual final position in the phrase to some other place within the phrase is also possible.)

In the more recent literature, a number of authors have suggested that the intonation phrase, rather than a syntactic unit, is a domain of phrase-final lengthening (Cruttenden, 1986; Beckman & Edwards, 1990; Gussenhoven & Rietveld, 1992). The question whether melodic and temporal cues are separate and what their respective perceptual salience is remains open to debate (Wichmann & Knowles, 1995). However, listeners seem to agree to a large extent on where the boundaries of intonation phrases are. How they do so is a separate issue and for the present phonetic research a mystery to be unravelled.

Some analyses, particularly, some recent phonological studies of intonation, make a distinction between two types of phrases, intermediate and intonation phrases (Beckman & Pierrehumbert, 1986) or major phrases and tone groups (Ladd, 1986). I did not make any similar distinction in my study since, firstly, phrases which might be classified as intermediate phrases occurred only infrequently, and, secondly, Czech phonetic literature recognises only one 'type' of phrase - the intonation phrase.

2.5.2 Candidate domains: justification for their choice

The motivation for choosing the interpause stretch, the clause and the intonation phrase as candidate domains for articulation rate variation was the following.

Although it is not clear what kind of unit the interpause stretch is, whether it is a discourse unit or a planning unit of some sort (see section 1.3.1), its relation to pauses is a strong reason to consider this unit a plausible candidate for the domain of articulation rate variation. Pauses are an important part of the temporal

structure of speech and their relation to articulation rate was already discussed in section 1.3.3. Given this relation and also the fact that the role of pauses in speech planning has been demonstrated in the literature (e.g., Rochester, 1973; Butterworth, 1975, 1980b), it is possible to imagine that the speech planning during pauses involves also organization of articulation rate. Additionally, the interpause stretch is used in the investigation of intra-speaker variation of articulation rate most frequently but never with any justification (see the discussion in section 1.3.1). If the interpause stretch could be defined as a/the domain of articulation rate variation, this would provide the necessary justification for its use.

The clause was chosen to cover the syntactic level. The existence of a relationship between syntax and various aspects of timing and of prosody in general has been demonstrated in a number of studies, especially on English (e.g., Goldman-Eisler, 1972; Cooper & Paccia-Cooper, 1980; Selkirk, 1984). It is possible that the clause in Czech is a domain for similar phenomena too, and it might also function as a/the domain of articulation rate variation.

At the phonetic and phonological levels, the intonation phrase was the most obvious candidate for investigating articulation rate. It has been shown that the intonation phrase behaves not only as an intonational but also a rhythmical unit, at least in English and Czech (e.g., Boomer, 1965; Lehiste, 1970; Palková, 1994).

2.5.3 The unit for the measurement

There are various ways of measuring articulation rate within a unit such as the interpause stretch, the clause or the intonation phrase; in other words, there are various minimum units across which articulation rate can be measured. My choice was the phonological word. The thinking behind this goes as follows.

Starting from the smallest conceivable unit, options include: segment, syllable, morpheme, word, phonological word, syntagma. As I pointed out earlier, the criterion for the choice is theory-driven; it depends on our assumptions mainly about: (i) at what level of speech processing[7] articulation rate is planned, and (ii) for what unit or units articulation rate is planned.

My hypothesis is that articulation rate, as a prosodic feature, is likely to be planned at phonological or phonetic levels of speech production. This excludes morpheme, word and syntagma as candidates for the minimum unit of measuring articulation rate because these are grammatical units. As for the unit or units for which the articulation rate is planned, my assumption is that such a unit needs to be an element carrying a distinctive meaning. This rules out segments and syllables and leaves us with the phonological word.

The phonological word is defined as a string of syllables containing a single stress and respecting lexical word boundaries (Selkirk, 1980a,b; Grosjean & Gee, 1987; for specific characteristics of phonological words in Czech, see the following section). Apart from being a phonological unit carrying meaning, there are several other reasons why the phonological word is a suitable unit for measuring articulation rate variation within the interpause stretch, clause or intonation phrase. The first reason is connected with rhythm, the concept closely related to articulation rate. The phonological word is traditionally considered in the Czech phonetic literature (e.g., Ondráčková, 1954a,b; Palková, 1994) to be the smallest rhythmical unit, in the sense that speakers tend to divide their speech into phonological words of a similar number of syllables (predominantly two or three) and that the syllables have a similar duration. Secondly, there is some evidence for the word as a unit of tempo in the sense that accentuation seems to affect the duration of a single word rather than all words within a larger domain, e.g.,

[7] By 'speech processing' or 'speech production' I mean the whole process from conceptualisation to forming an articulatory (motor) plan (Levelt, 1989). This differs from, e.g., Kent, Adams & Turner (1996), who by 'speech production' understand only the phonetic stage and motor control; the other stages belong in their system to 'language formulation'.

constituents, and the word as a whole, all segments and syllables, are affected (Eefting, 1991). Thirdly, it is also the minimal unit which can form an intonation phrase (Crompton, 1980). Fourthly, some theories of speech production (e.g., Levelt, 1989) explicitly recognize the phonological word as an important unit at the phonological level of the processing of connected speech. Finally, a practical consideration: the word is a small enough unit to make it possible to observe the variation of articulation rate in reasonable detail (see the discussion in section 1.3.2).

2.5.4 The phonological word in Czech

In Czech, where the stress falls regularly on the first syllable, the phonological word usually coincides with a lexical word, e.g. (slashes stand for word boundaries):

ˈ *Jeho* / ˈ *soused* / ˈ *četl* / ˈ *celé* / ˈ *odpoledne* /

(His / neighbour / read [was reading] / the whole / afternoon.)

Sometimes the phonological word includes one or more enclitics (underlined; double slash stands for the intonation phrase boundary), e.g.:

ˈ *Když jsem to* / ˈ *slyšel* / ˈ *poprvé* // ˈ *nevěřil jsem tomu* //

(When [I] + auxiliary 'to be' 1st sg.+ it / heard / the first time //
I didn't believe it //)

Unstressed monosyllabic words can also be attached to the following stressed word; they become proclitic. This normally happens only at the beginning of the intonation phrase and it is the only case in which stress falls on the second syllable in the phonological word, e.g.:

A ˈ*proto* / ˈ*řekl* // *že* ˈ*zůstane* //

(And therefore / [he] said // that [he] would stay //)

2.5.5 Measurement technique

Speech samples were digitized at a resolution of 16 bits and a sampling rate of 11 kHz using the SignalyzeTM 2.03 signal analysis software on an Apple Macintosh Quadra 950. Speech files were also copied to a Sun Sparcstation LX for analysis using Entropics waves+ speech analysis software.

Phonological words were identified impressionistically by myself. Identification of phonological words in Czech is usually unproblematic. Most phonological words contain a single lexical word or a lexical word plus one or more enclitics, and the word boundary is normally just before another stressed syllable starts. There were a small number of cases where it was difficult to decide whether a function word was an enclitic or a separate phonological word. Very exceptionally a difficulty also was deciding whether a monosyllabic function word (it was always a reflexive particle 'se') is an enclitic or a proclitic. These cases were decided by comparing the solutions from another phonetician.

In order to make measurements of the duration of phonological words and pauses as accurate as possible, a combination of impressionistic listening and reference to waveform displays and wideband spectrograms was used.

Identifying the boundaries of interpause stretches was straightforward; they were determined by finding pauses 130 ms or longer. Clause boundaries were also unproblematic to determine in the text analysis. Intonation phrases were identified impressionistically by myself, and, independently, by another Czech phonetician to avoid idiosyncrasy. Perceptually, the intonation phrase is a salient unit; it has been claimed that listeners tend to identify these phrases with considerable

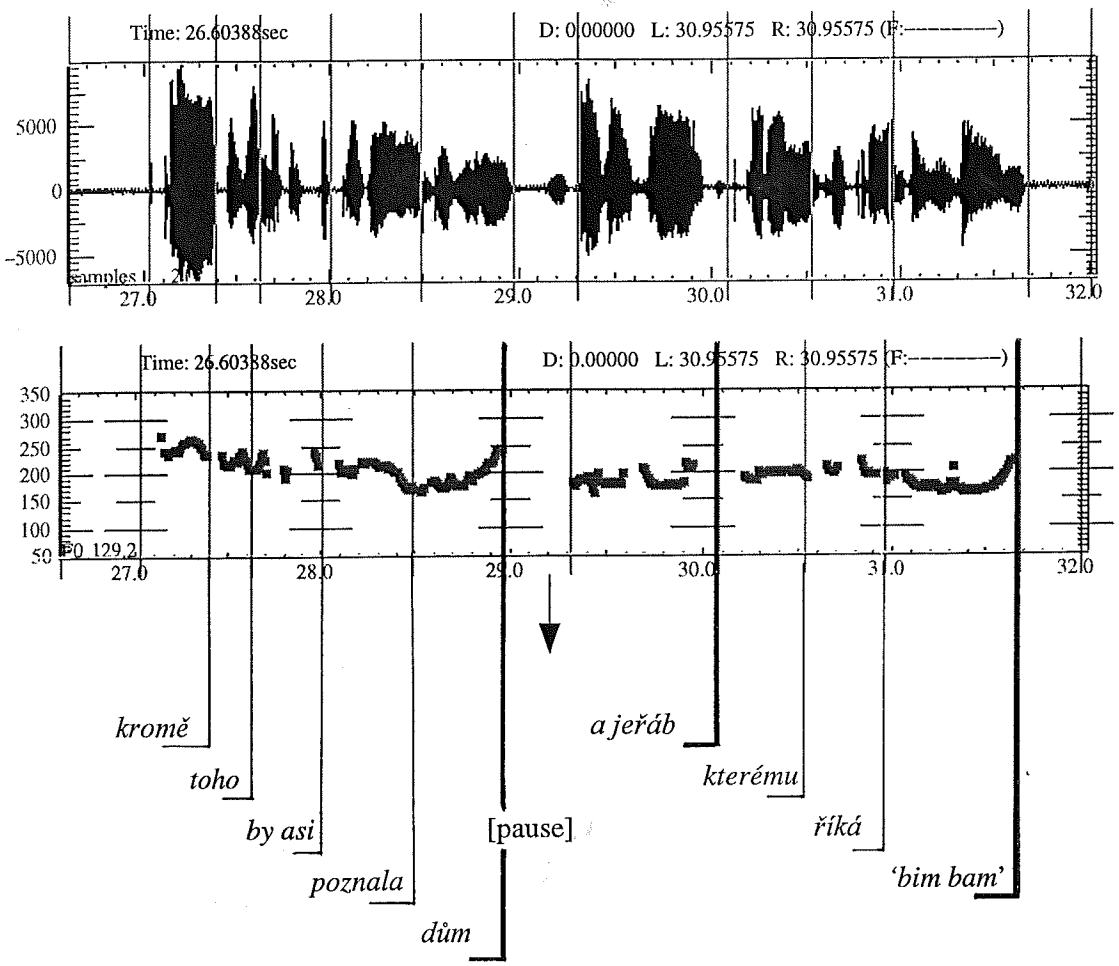
agreement (e.g., Palková, 1994). The agreement between my analysis and the other phonetician's analysis was very high. The few cases of disagreement were then adjudicated by a third phonetician.

Figure 2.1 presents a waveform and a f_0 track (obtained using the ESPS 'formant' routine) of three intonation phrases from one speaker, with vertical lines indicating the measurement points. Thin lines represent word boundaries, thicker lines intonation phrase boundaries. The three intonation phrases are (between double slashes):

// *kromě toho by asi poznala dům* // (pause) *a jeřáb* // *kterému říká 'bim bam'*//

// apart from that she would perhaps recognize a house // and a crane // which she calls 'bim bam' //

Figure 2.1 **A waveform and a f_0 track of three intonation phrases from speaker MS, with vertical lines indicating the measurement points. Thin lines represent phonological word boundaries, thicker lines intonation phrase boundaries.**



2.5.6 The unit for expressing articulation rate

In this thesis, articulation rate was expressed in terms of the number of syllables per second. The other possible unit - the segment - was disregarded for the following reasons:

(i) The first reason is practical. For the purpose of this study, the question is how many syllables or segments there are in the phonological word. Intuitively, people can count syllables without major difficulties. In counting segments, the

situation is less straightforward. The segment is a theoretical construct, and thus it depends on our theory whether, for instance, we count affricates or phonemically long vowels (which Czech has) as one or two segments. Defining syllable boundaries and structure may be problematic, however, that does not concern this study.

(ii) The syllable is sometimes believed to be the minimum planning unit (e.g., Crompton, 1981)[8] and the smallest unit of articulation. Evidence on coarticulation within a syllable can be found in a number of studies (e.g., Kozhevnikov & Chistovich, 1965; Liberman, 1970; Whalen, 1990). Further evidence is provided in the research on contextual effects on vowel duration (e.g., van Santen, 1992) or on allophonic variations of consonants (e.g., Kahn, 1976; Church, 1983), the position within a syllable being crucial. The role of the segment in speech production is more controversial than the role of the syllable.

(iii) The syllable in Czech is also considered a minimum rhythmical unit. Traditionally, Czech is claimed to be a syllable-timed language (Palková, 1994). However, some authors prefer to characterize it as having both syllable- and stress-timing (e.g., Daneš, 1957; Bolinger, 1965). It must be noted, that the basis for distinguishing between syllable-timing and stress-timing is highly controversial (e.g., Roach, 1982; Dauer, 1983; Williams & Hiller, 1994).

(iv) Most studies on articulation rate use nowadays the measure of the number of syllables per second. To be able to compare the results with findings about other languages, use of the same measure is necessary. The consequences of comparing the articulation rate of two languages expressed in different measures were discussed in section 1.3.6.

---

[8] Evidence for this comes, for instance, from studies on language acquisition (e.g., Ingram, 1989).

For the purpose of this study, the syllable is defined in terms of the presence of a syllable nucleus (apart from vowels, syllabic /r/ or /l/ are syllable nuclei in Czech). The number of syllables in a phonological word was determined impressionistically by myself.

When examining articulation rate, a question often arises whether to count elements which disappear in speech, i.e., whether to consider the 'underlying form' of the word or its 'surface form'. In continuous speech, particularly in spontaneous speech, segments disappear quite frequently and it can happen to syllables as well. This was, however, not the case in the present study; all three speakers used standard, rather than colloquial Czech (where segmentally simplified versions of certain words are even lexicalized), and moreover, they were rather precise in their pronunciation.

2.6 DETERMINANTS OF ARTICULATION RATE VARIATION

As I outlined at the beginning of this chapter, I was interested not only in how articulation rate varies in speech, but also what factors determine its variation.

2.6.1 Position of the phonological word within the domain

In the part of the study concerned with statistical analysis, the position of each phonological word within a candidate domain was coded, and an analysis was carried out to find out whether the word's position is a factor which affects significantly the articulation rate variation. Significant results would be expected in case of a regular patterning of articulation rate variation.

Oxford University
Phonetics Laboratory

2.6.2 Size of the phonological word

For the analysis of the relationship between the size of the phonological word and its articulation rate, a decision had to be made as to which unit should be used to express the size of words. A common measure is the number of segments. An alternative is the number of syllables. In order to see which of the two is a better measure, both were examined in this analysis.

As I noted above, the segment is a theoretical construct, and thus it needs to be specified here whether phonologically long vowels and affricates count as one or two segments. In my case, the criterion was the reported average duration of these sounds in comparison with the duration of phonologically short vowels and, for affricates, a comparison with other consonants.

It has repeatedly been claimed in the Czech phonetic literature that long vowels are approximately twice as long as their short counterparts in the same position in the word (e.g., Chlumsk", 1928; Romportl, 1973). Thus it seems reasonable to assume that the presence of long vowels in a word may influence the duration of the word and of its constituent syllables and hence its articulation rate. In order to take this factor into account, two alternative procedures were used in counting segments in words: in one case, long vowels were counted as two segments (Palková & Ptáček, 1995), in the other, as one segment. The duration of the two (voiceless) affricates in Czech is not significantly longer than the duration of voiceless plosives or voiceless fricatives (Petr, 1986) and thus affricates were counted as one segment.

2.7 STATISTICAL PROCESSING

The main statistical processing was performed using Generalized Linear Modelling (GLM): General Factorial. This procedure provides analysis of

variance for one dependent variable by one or more explanatory variables. It is important that this method is able to test unbalanced as well as balanced data sets since, in my case, the number of observations (words) in various subsets of data was unequal. The GLM analysis was combined with a post-hoc Tukey test. If an overall F test shows that significant differences exist among the means, Tukey's honestly significant difference test (Tukey's HSD) can be used to determine, by using pairwise multiple comparisons, which means differ from each other at a statistically significant level.

In the present study, the dependent variable was articulation rate and the explanatory variables were: position of the word within the intonation phrase, the size of the word in syllables or segments, and speaker. All the explanatory variables were treated in the model as factors although position and word size are numerical variables and therefore they would normally be included as covariates. This was necessary since in the GLM analysis it is not possible to conduct a post-hoc Tukey test on covariates and this test was crucial for the purpose of the study.

In order to obtain an estimation of the power of the model in accounting for the variance in articulation rate, the coefficient of determination (R^2) was also examined. This coefficient is the proportion of the variation in the dependent variable determined by the variation in the factors (Ferguson & Takane, 1989).

The significance level used in Study 1 is $p < 0.05$.

The analysis had to be executed separately on groups of 2-word, 3-word and 4-word intonation phrases, given the coding of the factor position (for instance, position 2 in 2-word intonation phrases is not equivalent to position 2 in 3- or 4-word phrases).

All the statistical processing was carried out using SPSS for Windows, version 7.5.1.

CHAPTER 3

STUDY 1: RESULTS

The analysis of the intonation phrase will be described first, followed by the analyses of interpause stretches and clauses.

3.1 FREQUENCY DISTRIBUTION OF INTONATION PHRASES OF DIFFERENT SIZES

In total, there were 178 intonation phrases (henceforth, just 'phrases' or IP's) in the material. Table 3.1 shows, for each speaker, the distribution of the phrases according to the number of words they contained and the proportion of the total number of phrases for phrases of different size. The table also gives the means and standard deviations (SD) for phrases of different size. The total number of phonological words measured was 446.

Table 3.1 **Frequency distribution of intonation phrases according to their size.**

Speaker	Categories of IP's (Number of words in the IP)						Total no. of IP's
	1	2	3	4	5	6	
MS	13	8	12	7	10	1	51
OL	24	14	14	7	1	0	60
JP	18	18	15	11	3	2	67
Total	55	40	41	25	14	3	178
	31%	22%	23%	14%	8%	2%	
Mean AR	4.67	5.74	5.83	6.31	6.49	6.57	
SD [syll/s]	1.40	2.01	1.97	2.13	1.95	1.47	

As seen from Table 3.1, intonation phrases containing only one phonological word form the largest group. Intonation phrases of two and three phonological words are also frequent and after this point, the longer the intonation phrase, the less frequent is its occurrence (with exception of 5-phonological-word long intonation phrases in speaker MS). No intonation phrase was longer than six phonological words.

3.2 IS ARTICULATION RATE CONSTANT WITHIN THE INTONATION PHRASE?

Only 3-word, 4-word and 5-word phrases were considered in this part of the analysis since for shorter units, the term 'unit-internal variation' is not really appropriate, and for longer units, the number of examples in the database was very small. Two 3-word phrases were excluded from the analysis because they contained slips of the tongue (Table 3.1 does not include them).

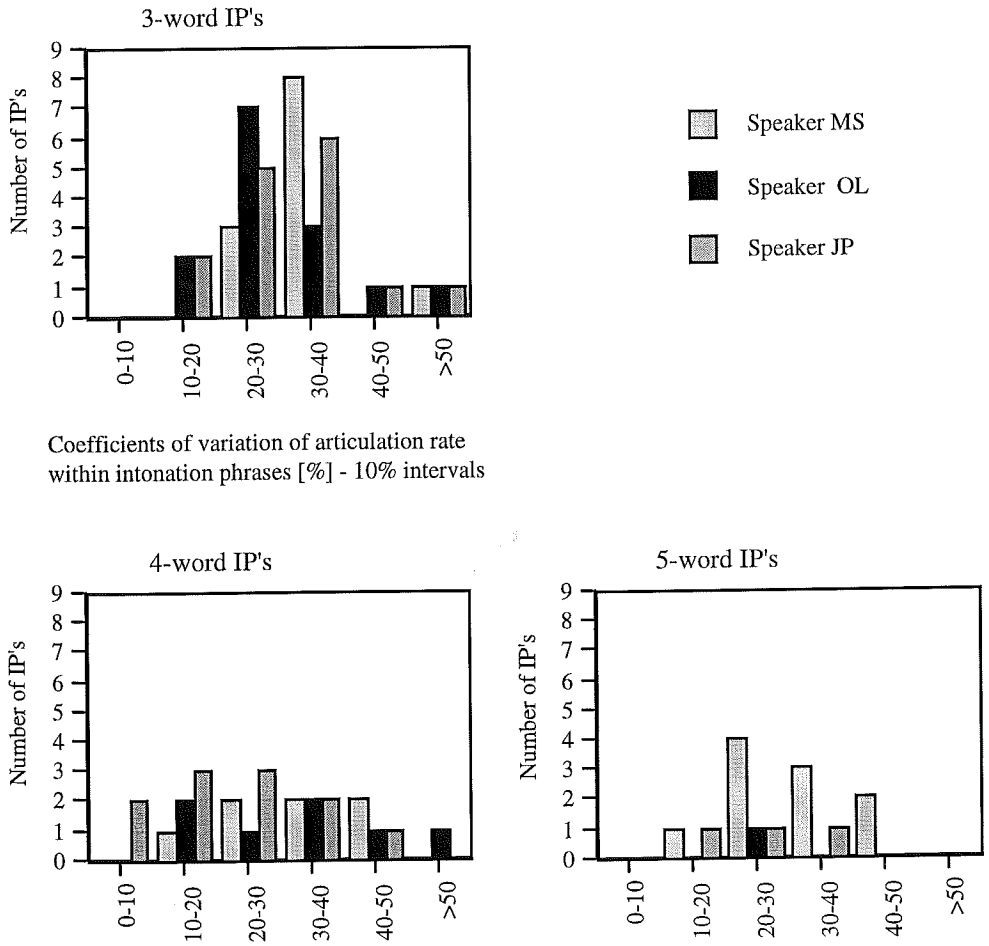
For testing the null hypothesis, i.e., whether articulation rate is constant within the phrase, the statistical measure of the coefficient of variation (CV) was employed. This is a common measure in studies on articulation rate. In general, it expresses variation in terms of a percentage of deviation from the mean. The formula is as follows:

$$CV = \frac{SD}{Mean} \times 100$$

Figure 3.1 shows, for each speaker, the extent of variation within phrases of different sizes. Articulation rate within phrases was far from constant: the overall range of variation was between 6.7% and 56.5% of the mean articulation rate within a phrase. In the 3-word and 5-word phrases, the coefficient of variation for the majority of phrases fell between 20% and 40% in all three speech samples. There was no apparent prevailing range in the variation in articulation rate of the 4-word phrases; the range was more widespread (10% - 50%). The smallest

variation (below 10%) was found only in two phrases in the sample from JP. Thus, overall, values below 10% can be described as exceptional. Variation above 50% is also very rare.

Figure 3.1 **Frequency distribution of intonation phrases according to their internal articulation rate variation. Columns represent the number of phrases for each of the 10% intervals of coefficient of variation.**



A separate statistical analysis for checking cross-speaker consistency in the degree of articulation rate variation was not conducted for two reasons. Firstly, the number of observations is too small; secondly, it will be demonstrated in Study 2

that when articulation rate of individual phonological words is compared across different speakers speaking the same text, it is remarkably consistent.

Evaluating the degree of variation in the present study in light of previous studies is not straightforward because there does not seem to be an established criterion for what is regarded as 'small' or 'considerable' variation in articulation rate. However, if our judgement is based on the conclusions (unfortunately, typically lacking in theoretical motivation or discussion) of previous studies, a threshold around 20% can be discerned (e.g., Goldman-Eisler, 1968; Grosjean & Deschamps, 1975; Miller *et al.*, 1984). This threshold seems, however, rather arbitrary. A missing reference point might be provided, for instance, by a perceptual threshold, i.e., just noticeable differences (JND) in articulation rate. This issue does not seem to have been investigated sufficiently; the only figure available for JND in articulation rate seems to be by Eefting & Rietveld (1989), who examined Dutch sentences and suggested a threshold of 4.43%. In comparison with the conclusions of previous studies on variation in articulation rate and, especially in comparison with JND, the articulation rate within the phrases in my material manifests considerable variation and thus the null hypothesis that articulation rate is constant within the phrase was rejected.

3.3 PATTERNING OF ARTICULATION RATE VARIATION WITHIN INTONATION PHRASES

The second hypothesis tested was that articulation rate variation within the phrase follows a certain recurrent pattern or patterns. This part of the analysis was based first on the rank ordering of component words according to their articulation rate; absolute values were not considered at this stage. This allowed the identification of any ordinal patterns of articulation rate variation which might form a typical structuring of variation within phrases. Two-, three- and four-word phrases were examined. Five-word phrases were not considered since the number of tokens was

very uneven across speakers, and in two speakers too small to allow any tendencies to become apparent (see Table 3.1).

The results are shown in Tables 3.2-3.4. The patterns of articulation rate variation are represented iconically and, for later convenience, by a sequence of letters. In the iconic representation, dots stand for words, falling or rising lines between the dots indicate the direction of change in the articulation rate. Thus higher dots represent a faster articulation rate than lower dots. In letter representation, each letter stands for a single word, in such a way that 'A' represents the first word in a phrase; 'B' the second, etc. Letters (words) are arranged in sequences in descending order of their articulation rate. For instance, the sequence BAC indicates that the second word in the phrase was the fastest, the first was slower and the last was the slowest.

Table 3.2 shows that in 85% of 2-word phrases, the second word was pronounced more slowly than the first one (pattern AB).

Table 3.2 Ordinal patterns for 2-word intonation phrases.

	AB 	BA 
Speaker		
MS	6	2
OL	14	-
JP	14	4
Total	34	6

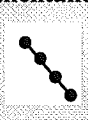
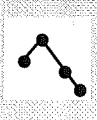
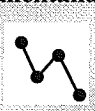
Table 3.3 Ordinal patterns for 3-word intonation phrases.

	ABC 	BAC 	BCA 	ACB 	CAB 
Speaker					
MS	7	4	-	-	1
OL	5	6	2	1	-
JP	8	6	1	-	-
Total	20	16	3	1	1

In the case of 3-word phrases (Table 3.3), the prevailing patterns of articulation rate were ABC, i.e., decreasing articulation rate throughout the phrase (49% of the total number of phrases), and BAC, where there is an increase in rate first and then decrease towards the end of the phrase (39%). The frequency of these two patterns in the individual speakers was fairly balanced. In 98% of all 3-word phrases the fastest word was the first or the second one; there was only one phrase in which it was the last word that had the highest articulation rate in the phrase. In fact, the last word was usually the slowest (88% of cases).

In 4-word phrases (Table 3.4), the trend was similar to that in 3-word phrases. The first or the second word tended to be the fastest in the phrase (40% of the total number of phrases in each case). In 20% of cases, the third word was fastest. There were no phrases in which the last word had the highest articulation rate. Again, the last word tended to be the slowest (68% of phrases).

Table 3.4 Ordinal patterns for 4-word intonation phrases.

	ABCD rallentando 	**BACD rallentando** 	**BCAD rallentando** 	**ACBD rallentando** 
Speaker				
MS	3	2	-	-
OL	3	1	1	-
JP	1	2	1	1
Total	7	5	2	1

	BCDA rallentando 	**ADBC rallentando** 	**ACDB rallentando** 	**CABD rallentando** 
Speaker				
MS	-	1	-	-
OL	-	-	1	-
JP	1	-	-	1
Total	1	1	1	1

	CBDA accelerando	CBAD accelerando	CADB	BADC
				
Speaker				
MS	-	-	1	-
OL	-	1	-	-
JP	1	-	1	2
Total	1	1	2	2

As far as the overall structure of articulation rate within 4-word phrases is concerned, twelve different patterns occurred. However, it seems quite obvious that not all the patterns are equally distinct. There are a number of patterns which are alike in the sense that at least three out of four component words follow the same direction of articulation rate change. In particular, two distinct subgroups of patterns can be recognised - rallentando (progressive decrease in articulation rate) and accelerando (progressive increase in articulation rate). According to this criterion, the first eight patterns (ABCD, BACD, BCAD, ACBD, BCDA, ADBC, ACDB, and CABD) can be classed as rallentando, covering 76% of 4-word phrases. Two patterns (CBDA and CBAD) involve a progressive increase in the articulation rate - accelerando (8% of phrases). The remaining two patterns (CADB and BADC), found in 16% of phrases, are not sufficiently similar to one another to permit a single criterion to characterize them both.

When the samples were considered individually for their 4-word phrases, speaker JP stands out as different from MS and OL by having a relatively balanced number of phrases with rallentando as opposed to non-rallentando patterns. The validity of this observation was tested in the next section, based on the absolute values of articulation rate.

3.4 INTERNAL STRUCTURE OF ARTICULATION RATE VARIATION WITHIN THE INTONATION PHRASE

The analysis based on ordinal patterns suggested that articulation rate does reflect a systematic pattern within intonation phrases and that the normal pattern involves a progressive slowing down (rallentando). The main aim of further analysis was to test the statistical significance of this slowing down.

Hypothesis 4.1:
Articulation rate variation within the intonation phrase reflects a regular pattern of rallentando (progressive slowing down) in all three speakers.

Hypothesis 4.1 was tested by the GLM analysis whose nature was described in more detail in section 2.7. From a statistical point of view, testing this hypothesis amounts to testing whether the factor of position of the word within the phrase has a significant effect on its articulation rate. The lack of a significant effect would indicate that the mean articulation rates of words within the intonation phrase are not statistically different from each other. However, phrase-final lengthening is likely to contribute to the rallentando pattern. If its effect on word duration as a whole is strong enough and consistent, this phenomenon on its own could lead to an overall significant effect.

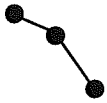
How the articulation rate decreases with a later position within the phrase becomes obvious from examining the mean rates in different positions. There are two basic possibilities, or shapes, for this decrease. The first one involves an even spread of the slowing down throughout the phrase; it is linear. In the second possibility, the spread is not even; it is non-linear and it can have several 'subshapes'. The schematic pictures below will help to visualize the two scenarios for phrases containing three words.

(1) Linear slowing down

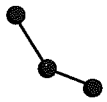


(2) Non-linear slowing down

a)



b)



Four word phrases would work on the same principle, with more possible 'subshapes' under (2).

As will be shown below, the GLM analysis confirmed hypothesis 4.1. The next step was to test whether the mean articulation rates of *all* words within the phrase are statistically different. In general, for the linear slowing down, this would represent only a checking procedure since the significant result of the GLM analysis, together with even differences between the mean rates between all the positions within the phrase, would already predict that all the differences are significant. In the case of a non-linear slowing down, however, examining the differences between the mean rates in different positions is crucial. For instance, in 2a the concern would be whether the mean articulation rates of the first two words are statistically different. These considerations are formulated here as hypothesis 5.1:

> **Hypothesis 5.1:**
> **The articulation rates of words in all positions within the intonation phrase are significantly different from each other in all three speakers.**

The post-hoc Tukey test was used to test this possibility.

As described in sections 2.1 and 2.6, apart from the factor of position, the factor of the size of the word was considered in the present study as a factor which also

might significantly affect the articulation rate of the word. The hypothesis (introduced in 2.1 as hypothesis 3.1) is for convenience repeated here:

> **Hypothesis 3.1:**
> **The size of the phonological word significantly affects the word's articulation rate in all three speakers.**

The factor of speaker was also considered as a possible source of significant changes in articulation rate since it has been suggested in the literature that people differ in their articulation rate (see the discussion in section 1.3.2). The hypothesis is as follows.

> **Hypothesis 6.1:**
> **Speakers differ significantly from each other in their articulation rate.**

Hypotheses 3.1-6.1 were tested simultaneously, by including the factors position, word size and speaker in the GLM analysis and running a series of post-hoc Tukey tests.

3.4.1 Assumptions for GLM analysis

There are three basic assumptions underlying the use of the analysis of variance. The first one is that the data are a random sample from a normal population; the second one is that the variances of observations are equal across groups; the last assumption is that observations are statistically independent given the group to which they belong.

Tests of normality of distribution and homogeneity of variance were carried out for the present material. The Kolmogorov-Smirnov test suggested that in some

cases the assumption of the normality of distribution was not satisfied. However, the Kolmogorov-Smirnov test is a sensitive test and it can be affected even by a small number of outliers while the distribution is, in fact, normal. Examination of Normal Q-Q plots showed that this must be, indeed, the case in the present material; the plots did not show any massive departures from normality. Given this further test and also the fact that the analysis of variance is robust to departures from normality, this assumption can be considered unproblematic for the present material.

As to the second assumption, the Levene test demonstrated that the variance was homogeneous across levels within each factor.

Finally, the most important assumption - the independence of observations one from another, can be considered fulfilled given the nature of the experimental design; the subjects produced continuous speech (no repetitions) under the same conditions. There was nothing in the experimental setup which would provoke subjects to make the articulation rate of one word influence the articulation rate of another. However, it is conceivable that physiological constraints on articulation or speech planning might undermine the claim about the independence of observations. But given the lack of any clear evidence for such constraints I shall proceed in my analysis assuming the independence of observations.

3.4.2 Effect of the position of the word within the intonation phrase and its size on the word's articulation rate: An overall model

An overall general linear model was designed, based on the data from all three speakers together, but, as explained in section 2.7, for the groups of 2-, 3- and 4-

word phrases separately. The three factors[1] were speaker, position and word size, the dependent variable was the word's articulation rate. This model was used to test the relationship between articulation rate of the word, and its position and size, and whether there were significant differences between speakers. Since word size was expressed in three different ways (in number of syllables, in number of segments while counting long vowels as one segment, and in number of segments while counting long vowels as two segments), the GLM analysis was run three times.

In all cases, a saturated model was used first, to find out whether there were any significant interactions. In general, if interactions are non-significant, it means that the factors are independent from each other. The interaction of the greatest interest here was position*speaker, which, if significant, would indicate differences between speakers as to the patterns of articulation rate variation within the intonation phrase. The other 2-way interactions, position*word size and speaker*word size were predicted to be non-significant since it is unlikely that the trend for the articulation rate across words of different sizes would be greatly dissimilar across speakers or in different positions within the intonation phrase.

The analysis, of which the full results are presented in Appendix 1, showed the following. The interaction speaker*position was not significant in any of the groups of phrases, indicating that the speakers were not statistically different in their patterns of articulation rate. This is crucial especially for 4-word phrases, since in the analysis of ordinal patterns, speaker JP seemed to differ from the other two speakers by having a relatively balanced number of rallentando and non-rallentando patterns. Yet, the statistical analysis showed that this speaker did not, in fact, differ in her articulation rate patterns from the other two speakers. Other 2-way interactions, position*word size, speaker*word size, and the 3-way

[1] Although the explanatory variables position and word size are numerical, not nominal variables, they were used as factors and not covariates at this stage of the GLM analysis since otherwise it would not be possible to carry out a Tukey test.

interaction position*size*speaker were also non-significant ($p > 0.05$). Subsequently, a GLM analysis was run only with the main factors. It yielded the following results (in full, they are included in Appendix 2).

The size of the word was significant in all groups of phrases but only when expressed in terms of the number of syllables; for 2-word phrases [F (4,71) = 3.72, $p < 0.008$], for 3-word phrases [F (4,113) = 7.84, $p < 0.001$],[2] and for 4-word phrases [F (4,90) = 3.25, $p < 0.015$]. There was a clear tendency for the articulation rate to increase with increasing number of syllables in the word. However, a Tukey test showed that not all words were significantly different from each other in their articulation rate. In fact, in 3-word phrases, only the words containing either one or five syllables were significantly different from the other size words; in 4-word phrases, it was only the monosyllables which were significantly different from all the other words. In 2-words phrases, the picture is less clear; the only words which were significantly different in the articulation rate were the monosyllables vs. the 5-syllable words, and the disyllables vs. the 3-syllable and 5-syllable words.

When word size was expressed in terms of the number of segments (whether long vowels counted as two segments or one), no significant relationship between the word's articulation rate and its size was found. I shall return to this finding later in this chapter.

Since the only significant way of expressing word size turned out to be in terms of syllables, the results for the other two factors, speaker and position, reported below, are based on the models with syllables only. The factor of speaker was significant in 4-word phrases; [F (2,90) = 8.43, $p < 0.001$]. However, a post-hoc Tukey test revealed that only speakers OL and JP were significantly different in their mean articulation rates (speaker JP being faster, with 7.07 syll/s than speaker

[2] In both 2-word and 3-word intonation phrases, the word with six syllables was excluded from the analysis, since the Tukey test cannot be performed unless each of the levels within the category has more than one observation.

OL, with 5.15 syll/s; speaker MS's mean articulation rate was 6.26 syll/s). The factor of speaker was also significant in 3-word phrases, although rather weakly; [F (2,113) = 3.65, $p < 0.029$]. In fact, a Tukey test showed that none of the speakers was significantly different in the articulation rate from the others.

The factor of position was significant in all three groups of phrases; for 2-word phrases, [F (1,71) = 24.73, $p < 0.001$]; for 3-word phrases, [F (2,113) = 27.34, $p < 0.001$], and for 4-word phrases, [F (3,90) = 11.91, $p < 0.001$]. Further, the patterns of articulation rate were analysed by comparing mean articulation rates in different positions within the phrase and conducting a post-hoc Tukey test on the means. The results are summarized in Table 3.5.

Table 3.5 **Mean articulation rates [sylls/s] (summed across speakers) and standard deviations for words in different positions within the intonation phrase, and the results of the Tukey test ($p < 0.05$) on which positions involved significantly different means. (Positions are represented by numbers, e.g., 1-2 indicates that the mean articulation rates in the positions 1 and 2 within the IP are significantly different.)**

	2-word IP's		3-word IP's		4-word IP's	
	Mean	SD	Mean	SD	Mean	SD
Position 1	6.70	1.86	6.62	1.80	7.27	1.91
Position 2	4.72	1.59	6.62	1.88	6.98	2.42
Position 3			4.28	1.21	6.46	1.50
Position 4					4.52	1.47
Significantly different means (positions)	1-2		1-3 2-3		1-4 2-4 3-4	

The overall patterns of articulation rate were, for samples from all speakers together, AB for 2-word phrases, ABCD for 4-word phrases and A=BC for 3-word phrases (A=B indicates that words in the first two positions had identical mean articulation rates). These results generally confirm the observation made on the basis of ordinal patterns (section 3.3) that articulation rate within the phrase

tends to follow a pattern of slowing down. However, the values for the mean articulation rates in Table 3.5 indicate that the slowing down was non-linear and resembled schematic pattern 2a in section 3.4. A Tukey test showed that the first two words in 3-word phrases and the first three words in 4-word phrases were not significantly different from each other in their mean articulation rates. They were, however, significantly different from the last word within the phrase. A possible explanation of this is the occurrence of phrase-final lengthening.

The coefficient of determination (R^2) showed that in 2-word phrases, the model with the factors position, speaker, and word size in syllables, accounted for 41.5% of variance in articulation rate; in 3-word phrases, for 48.2%, and in 4-word phrases, for 47.3% of variance.

The equal means for the first two words in 3-word phrases are not surprising given that the most common patterns were ABC and BAC, which are opposite in the direction of the rate change in these two words and balanced in the number of occurrences. In order to find out whether equal means for the first two words apply to all three speakers, I examined patterns of articulation rate variation for individual speakers. Table 3.6 presents the results.

Table 3.6 **Mean articulation rates [sylls/s] and standard deviations (SD) for words in different positions within the intonation phrase: the results for individual speakers in 3-word phrases. The results of the Tukey tests ($p < 0.05$) on which positions involved significantly different means.**

	MS		OL		JP	
	mean	SD	mean	SD	mean	SD
Position 1	6.85	1.25	6.03	1.80	6.94	2.08
Position 2	5.98	1.50	6.46	1.83	7.27	2.11
Position 3	4.06	1.38	4.17	1.16	4.57	1.14
Significantly different means (positions)	1-3 2-3		1-3 2-3		1-3 2-3	

The table shows that the first two words do not have equal means in the individual speakers. The speakers differ in the overall patterns; subject MS gives an overall pattern ABC while speakers OL and JP pattern BAC. However, in none of the three speech samples were the first two words statistically different in their mean articulation rates. The word which stands out as being significantly different from both others is the last one. This might explain why in the saturated model, the interaction position*speaker was not significant in 3-word phrases and yet differences between patterns were found between speakers.

3.4.3 Patterns ABC and BAC

When looking at the ordinal patterns for 3-word phrases in section 3.3 and the results of the statistical analysis, which suggested that phrase-final lengthening operates within the intonation phrase, it is not a great surprise that it were the patterns ABC and BAC and not other patterns that occurred most frequently; only in these two patterns the final word is the slowest. However, no analysis so far allowed a direct assessment of the relationship of these two patterns. Thus I shall take a closer look at these patterns in this section.

As I mentioned above, the mean articulation rates of the first two words within the phrase were not statistically different either for speaker MS (pattern ABC) or for OL and JP (patterns BAC). However, the difference between the means of these words in samples OL and JP was, in fact, half as large as in sample MS (0.43 syll/s, 0.33 syll/s, and 0.87 syll/s respectively). The difference is small, but it suggests that, for 3-word phrases, pattern BAC may be, after all, weaker than pattern ABC. Further statistical analyses, identical to the ones presented in section 3.4.2, were conducted for patterns ABC and BAC separately to test their relative strength.

The results of the GLM analysis on position and size as the main factors and post-hoc Tukey test (Table 3.7) did support the claim that pattern ABC is stronger than pattern BAC. For both patterns ABC and BAC, the only significant factor was position. In ABC, the significance was strong for all three speakers; for MS, [F (2,18) = 20.97, p < 0.001], for OL [F (2,12) = 7.67, p < 0.007], and for JP, [F (2,21) = 13.86, p < 0.001]. In BAC, the factor of position was highly significant only for speaker MS, [F (2,9) = 12.65, p < 0.002], while for the other two speakers, OL and JP, the significance was weaker than in ABC; for OL, [F (2,15) = 5.82, p < 0.013]; for JP, [F (2,15) = 5.64, p < 0.015]. The full details can be found in Appendix 3.

Table 3.7 **Patterns ABC and BAC: Mean articulation rates [sylls/s] and standard deviations for words in different positions within the intonation phrase, and the results of the Tukey test on which positions involved significantly different means ($p < 0.05$). (Positions are represented by numbers, e.g., 1-2 indicates that the mean articulation rates in the positions 1 and 2 within the IP are significantly different.)**

Pattern ABC

	Speaker MS		Speaker OL		Speaker JP	
	Mean	SD	Mean	SD	Mean	SD
Position 1	7.45	1.12	7.15	1.46	8.24	1.04
Position 2	5.46	1.42	5.69	1.07	6.80	1.45
Position 3	3.69	0.52	4.21	0.98	4.96	1.22
Significantly different means (positions)	1-2 1-3 2-3		1-3		1-3 2-3	

Pattern BAC

	Speaker MS		Speaker OL		Speaker JP	
	Mean	SD	Mean	SD	Mean	SD
Position 1	6.18	0.98	5.53	1.74	5.82	1.93
Position 2	7.16	1.18	7.26	2.41	8.12	2.81
Position 3	3.70	0.82	3.65	1.12	4.16	1.00
Significantly different means (positions)	1-3 2-3		2-3		2-3	

The difference in strength between ABC and BAC was also reflected in that ABC involved a higher number of positions where significantly different mean articulation rates were found (Table 3.7). In ABC from speaker MS, all three positions were significantly different from each other while in her BAC pattern, only two pairs of mean articulation rates were statistically different. In ABC from JP, two pairs of means differed significantly while in her BAC pattern, only one pair did. For speaker OL, both patterns ABC and BAC involved only one pair of significantly different mean articulation rates.

Furthermore, for speakers OL and JP, the factor of position on its own accounted for a higher proportion of variance in articulation rate in ABC than it did in BAC (56.1% vs. 43.7% in OL and 56.9% vs. 42.9% in JP).[3] It is worth noting that OL and JP, the speakers who demonstrated a relative weakness of pattern BAC as opposed to ABC most clearly, were the same speakers who gave an overall pattern BAC when 3-word phrases were examined. This overall result was thus rather misleading since closer analysis suggests that pattern ABC is, in fact, stronger than BAC.

The next step in the analysis involved considering that there might be a reason for the particular distribution of patterns ABC and BAC in the domain of linguistic structure. A plausible general hypothesis might be that patterns ABC and BAC differ in terms of distribution of different sizes of word in different positions within the intonation phrase. Given that a tendency was found for the word's articulation rate to increase with the number of syllables in the word, it is possible that in pattern ABC we might find mostly polysyllabic words in the first position (position A), while in pattern BAC, they would be found in the second position (position B). However, this is unlikely since the factor of word size was not significant in the statistical analysis when each pattern was analysed separately.

[3] In MS, the scores for ABC and BAC were reversed, i.e., higher proportion of explained variance in the articulation rate applied to BAC. However, the difference is rather small (70% in ABC vs. 73.8% in BAC).

Figure 3.2 below shows the distribution of words according to their size in syllables in different positions within the phrase.

Figure 3.2 **Distribution of phonological words in different positions within the intonation phrase according to their size in syllables.**

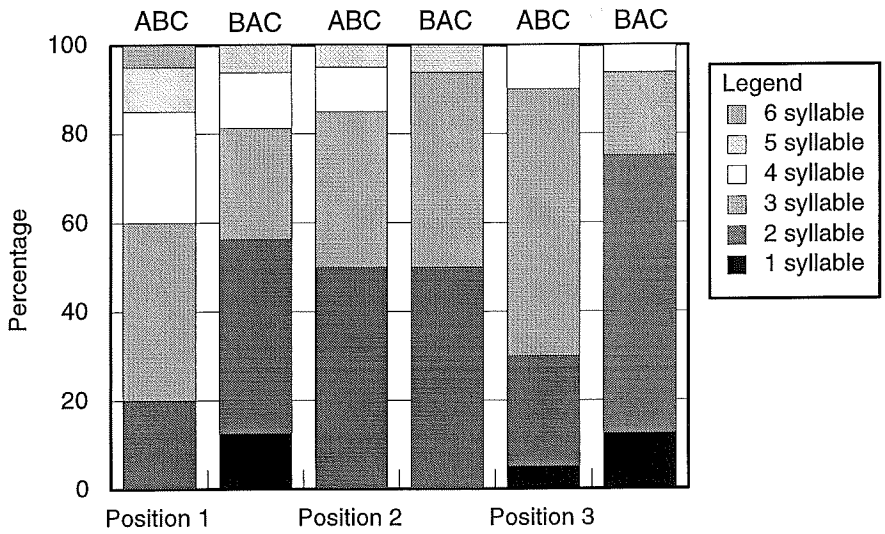


Figure 3.2 suggests that there is no clear bias in the distribution of phonological words of different size that could explain the occurrence of the BAC pattern. It is true that there is a tendency for the first position in the ABC pattern to have largely words with three or four syllables (65% together, as opposed to 38% in the BAC pattern), while in the BAC pattern, the first position involves a much larger percentage of monosyllabic and disyllabic words (57%, as opposed to 20% in the ABC pattern). This could contribute to some extent to a faster rate in the first position than in the second position in pattern ABC and vice versa in pattern BAC. However, in the second position of the BAC pattern, there is no tendency for polysyllabic words to prevail, nor for shorter words to prevail in the second position of the ABC pattern. In fact, the distribution of words according to their size is very similar for both patterns in this position. These observations were tested statistically by using a GLM analysis in which the dependent variable was articulation rate, and explanatory variables were position, word size in syllables,

speaker, pattern type and a 3-way interaction position*word size*pattern type. The 3-way interaction was of main interest, and it turned out to be non-significant, indicating that the difference in articulation rate variation in ABC and BAC cannot be explained in terms of the distribution of words of different size within the phrase. Thus there must be some other reasons for patterns ABC and BAC to occur with a similar frequency in speech. But since there is some evidence, however weak, for the ABC pattern being a stronger pattern, we could consider pattern ABC to be the default pattern, which is affected by a number of various factors causing deviations. More data is necessary and thus I shall return to this issue in Study 2.

Apart from pattern BAC there were a small number of other non-rallentando patterns in the material, as shown in Tables 3.3 and 3.4. I examined these patterns more closely, particularly with respect to their linguistic structure, discourse structure (repetitions, possible emphasis) and pauses surrounding some of the patterns. I was looking for regularities which could account for the shape of these patterns and thus could suggest additional factors for the GLM analysis. However, this investigation did not show any such regularities. This may be partly due to the very small number of observations.

3.4.4 Word size and articulation rate

As I mentioned earlier, a clear and overall significant relationship between the word's articulation rate and its size in syllables was found in my material. Surprisingly, no consistent trend was found when the size was expressed in segments. In this respect, it does not seem to be relevant whether long vowels count as two segments or just one.

Table 3.8 illustrates the results from the one-way ANOVA test, which was conducted separately for the factors of number of syllables and number of

segments, with articulation rate as the dependent variable in each case. The analysis was carried out for individual speakers (MS, OL and JP) and separately for 2-, 3- and 4-word intonation phrases. The table gives significance levels for explaining the variance in articulation rate (the full results for the cases in which significance was found are presented in Appendix 4). The data with an asterisk represent cases in which there was a clear tendency for the articulation rate to increase as the size of the word increased, although the tendency might not have been statistically significant. The results under 'number of segments' in the table are based on counting long vowels as one segment (given the results of the main analysis, no crucial difference was expected here between counting long vowels as one or two segments).

Table 3.8 **The levels of significance from one-way ANOVA for the factors of number of syllables and number of segments, taken separately, and with articulation rate as the dependent variable. For non-significant cases, $p > 0.1$. The asterisks indicate cases where there was a clear trend for the articulation rate to increase with an increasing number of syllables/segments.**

	Speaker	Number of syllables	Number of segments
2-word IP's	MS	0.003*	0.005*
	OL	ns	ns
	JP	ns	ns
3-word IP's	MS	ns*	ns
	OL	0.014*	ns
	JP	ns*	ns
4-word IP's	MS	ns	ns
	OL	0.001*	ns
	JP	ns*	ns

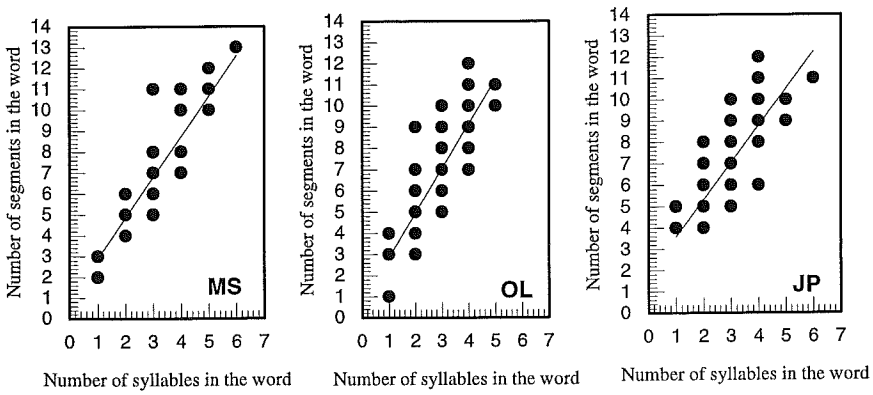
There could be two possible explanations why in most cases there was a clear tendency for the articulation rate to increase with the increasing number of syllables while there was no consistent trend for the number of segments: (i) if there is a tendency in Czech for multisyllabic words to contain syllables with a

simpler syllable structure, i.e., with a smaller number of consonants, than mono- or disyllabic words (F. Nolan, personal communication), and/or (ii) if there is some process of timing compensation within the syllable.

Although Czech might seem to foreign ears to have terrifyingly complex consonant clusters, the prevailing syllable type is CV, covering about 60% of tokens (Ludvíková, 1987). This does not rule out a possibility of some trading relationship between the syllable structure and the number of syllables in the word. To test this possibility, an additional analysis was carried out, firstly, for all the words grouped together within each speech sample, and, secondly, for the groups of intonation phrases separately within each individual speaker.

Initially, a correlation between the number of segments and the number of syllables in the word was examined. Since the results were consistent for both parts of the analysis, it will be sufficient to present only the findings for the individual speakers when all the words were grouped together. As Figure 3.3 shows, there is a strong positive linear correlation between the two. The Pearson correlation coefficients were as follows: speaker MS: $r = 0.895$ ($n = 80$), speaker OL: $r = 0.862$ ($n = 98$), and speaker JP: $r = 0.790$ ($n = 125$). All coefficients were significant at 0.1% level.

Figure 3.3 **Correlation between the number of segments and the number of syllables in the word.**



In other to investigate more directly the relationship between the number of syllables in the word and the syllable complexity, the average number of segments per syllable (effectively a measure of syllable complexity) for words containing a different number of syllables was analysed. For the calculation, long vowels were taken as one segment. Table 3.9 presents the results, together with the significance levels for the Pearson coefficient of correlation between the two variables. Since it is possible that the results could be affected by the cases with a very small number of observations, a correlation analysis was also conducted after excluding these cases.

The data in Table 3.9 indicate a clear tendency: the more syllables the word contains, the smaller the average number of segments per syllable is, in other words, the simpler the syllable structure. The Pearson coefficient of correlation revealed that this negative correlation is highly significant for all three speakers even after taking out the groups of words with a very small number of observations.[4] However, it should be pointed out that the maximum coefficient of determination (R^2) was 0.152 when comparing the three speakers and the two stages in calculation. This indicates that 15.2% of variance in word size is explained by the number of segments per syllable. This is a rather small effect. When the groups of 2-word, 3-word and 4-word intonation phrases were analysed separately within each speech sample, in all cases a negative correlation was found. However, only in 50% of cases was this tendency statistically significant.

Given the type and size of my data, it turned out to be difficult to carry out a rigorous statistical analysis for determining to what extent the significant correlation between the number of syllables in the word and its syllabic structure is responsible for the difference between the size in syllables and in segments in relation to articulation rate. As an indirect tool, I used the analysis of correlations

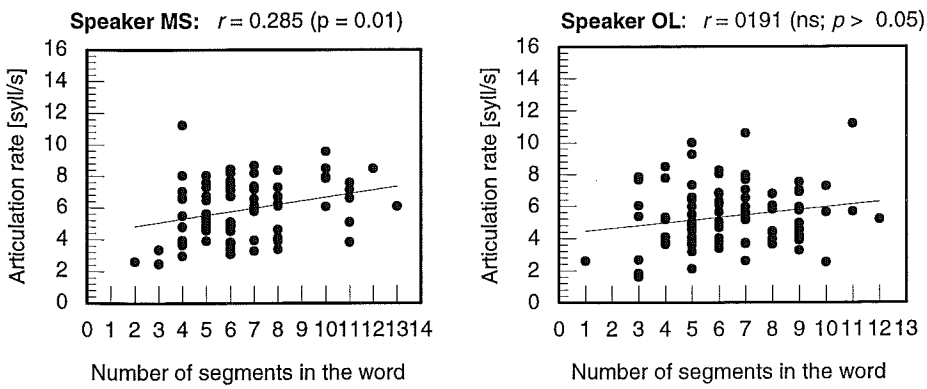
[4] The values for the Pearson coefficient of correlation are as follows; for the complete data sets, MS: $r = -0.305$, $p = 0.003$; OL: $r = -0.338$, $p < 0.001$; JP: $r = -0.391$, $p < 0.001$; for the data sets after excluding words with a small number of observations; MS: $r = -0.271$, $p = 0.009$; OL: $r = -0.323$, $p = 0.001$; JP: $r = -0.248$, $p = 0.003$.

between the word's articulation rate and, in turn, its size in segments, size in syllables and the number of segments per syllable. The results for individual speakers are presented in Figures 3.4, 3.5 and 3.6. Since the same trend was found when the groups of intonation phrases were examined separately as when they were grouped together, only the latter results are presented.

Table 3.9 **Average number of segments per syllable as a function of number of syllables in a word. The significance levels for the Pearson coefficient of correlation are given for the analysis before excluding words with a small number of observations (marked by an asterisk), followed by the significance levels after excluding them. Numbers in brackets represent the number of observations.**

The number of syllables in word	Speaker		
	MS	OL	JP
1	2.7 (3)*	3.2 (8)	4.5 (2)*
2	2.5 (38)	2.6 (38)	2.7 (60)
3	2.4 (36)	2.4 (38)	2.4 (42)
4	2.2 (12)	2.3 (12)	2.3 (17)
5	2.2 (5)	2.1 (2)*	1.9 (3)*
6	2.2 (1)*	-	1.8 (1)*
Significance of the correlation	$p = 0.003$ $p = 0.009$	$p < 0.001$ $p = 0.001$	$p < 0.001$ $p = 0.003$

Figure 3.4 **Correlation between the articulation rate and the number of segments in the word (2-, 3- and 4-word phrases combined).**



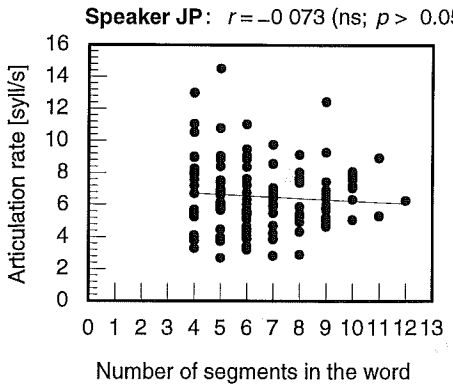
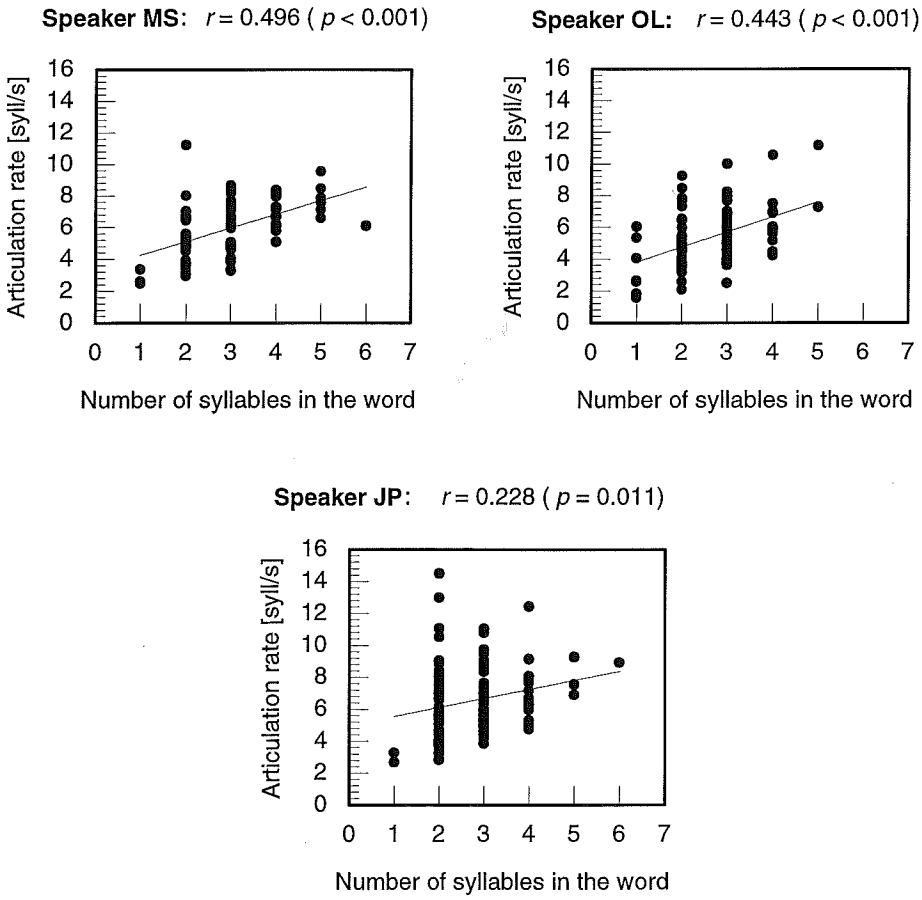
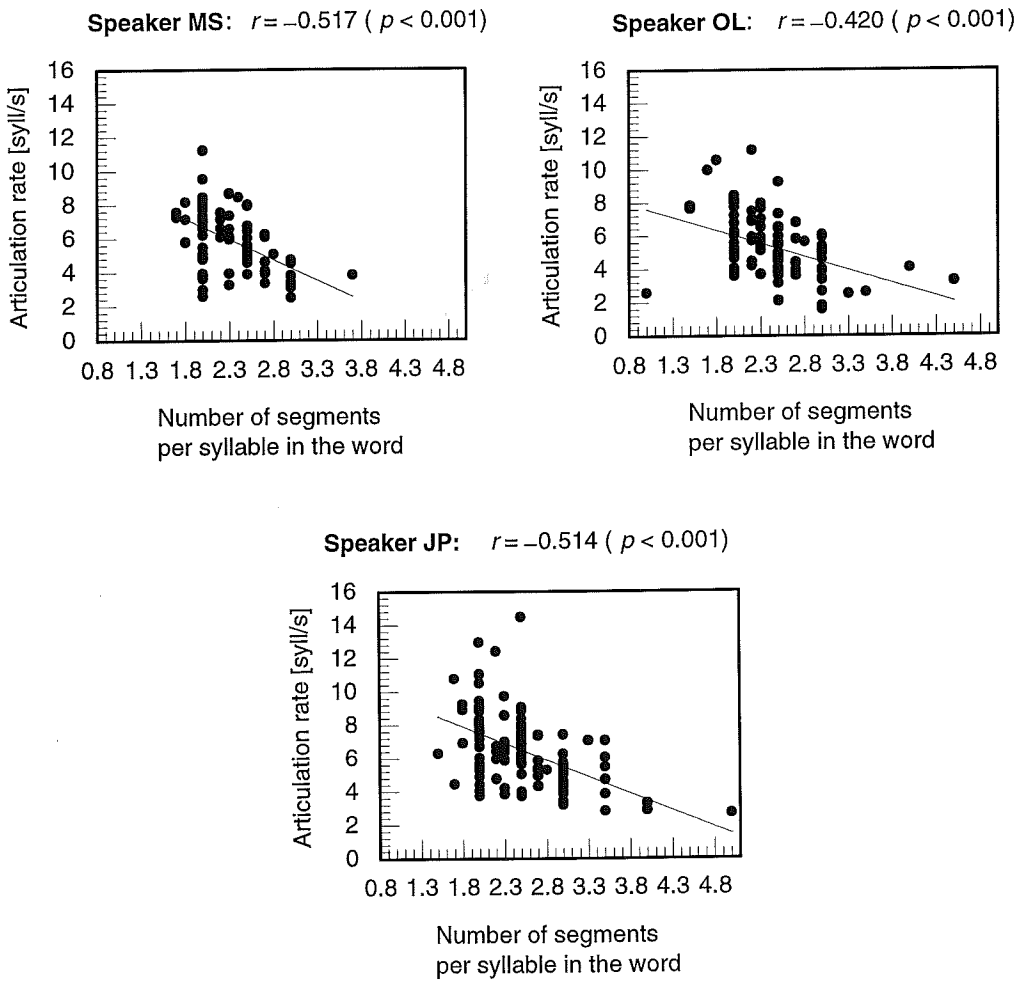


Figure 3.5 **Correlation between the articulation rate and the number of syllables in the word (2-, 3- and 4-word phrases combined).**



The significant positive correlation between the articulation rate of the word and its size in syllables is no surprise, given the findings from the ANOVA. As to the size in segments, there was no correlation, apart from one exception. Finally, a highly significant negative correlation was found between the number of segments per syllable and articulation rate. The positive correlation between articulation rate and the number of syllables together with the negative correlation between articulation rate and the number of segments per syllable may be taken as an explanation for the absence of correlation between articulation rate and the number of segments.

Figure 3.6 **Correlation between the articulation rate and the number of segments per syllable in the word (2-, 3- and 4-word phrases combined).**



3.5 INTERNAL STRUCTURE OF ARTICULATION RATE VARIATION WITHIN THE INTERPAUSE STRETCH AND THE CLAUSE

So far, I have been concerned in this chapter with the internal structure of articulation rate within the intonation phrase. The analysis showed that the intonation phrase fulfilled a criterion for being a domain of articulation rate variation, namely, it demonstrated a regular pattern of variation.
However, it is possible that interpause stretches and/or clauses either covary with intonation phrases or demonstrate different dominant patterns of variation. With respect to the relation between interpause stretches and intonation phrases, the situation that normally arises is that the interpause stretch boundaries coincide with the boundaries of intonation phrases; only exceptionally does a pause occur within a phrase. This means that interpause stretches are either coextensive with intonation phrases or contain a whole number of phrases (Daneš, 1957; Palková, 1994). Clauses also often coincide with phrases (Palková, 1994).

These generalisations also hold for the present material. No pauses occurred within phrases. Interpause stretches were coextensive with phrases in 66% of the total number of cases (119) and clauses in 59% of cases (for a total of 98 clauses). Conducting a separate analysis to test Hypothesis 1.1, that articulation rate is constant within the candidate domain, was pointless, given that in most cases boundaries of interpause stretches and clauses coincided with the boundary of an intonation phrase and a highly variable articulation rate was found within the intonation phrase.

The rate patterning found in these clauses and interpause stretches was, of course, identical to the patterning of the phrases they coincide with. It was thus more interesting to look at those interpause stretches and clauses which did not coincide with intonation phrases, i.e., contain more than one phrase. It is theoretically possible that the interpause stretch and/or clause are hierarchically higher domains of variation in articulation rate. A manifestation of this could be some kind of a

systematic patterning of articulation rate involving the whole interpause stretch or clause.

The analysis of the internal structure of articulation rate both within the interpause stretch and the clause was conducted in the same way as for the intonation phrase, i.e., using the GLM analysis. Since the analysis had to be performed separately for the units containing a different number of words and since there was an insufficient number of these per speaker, the data from all three speakers were pooled together. However, the possibility of differences between the speakers was tested by including 'speaker' as a factor into the model, alongside 'position' and 'size' in number of syllables or segments; articulation rate of phonological words was the dependent variable. The analysis was conducted on interpause stretches and clauses containing at least two phrases.

3.5.1 Interpause stretches

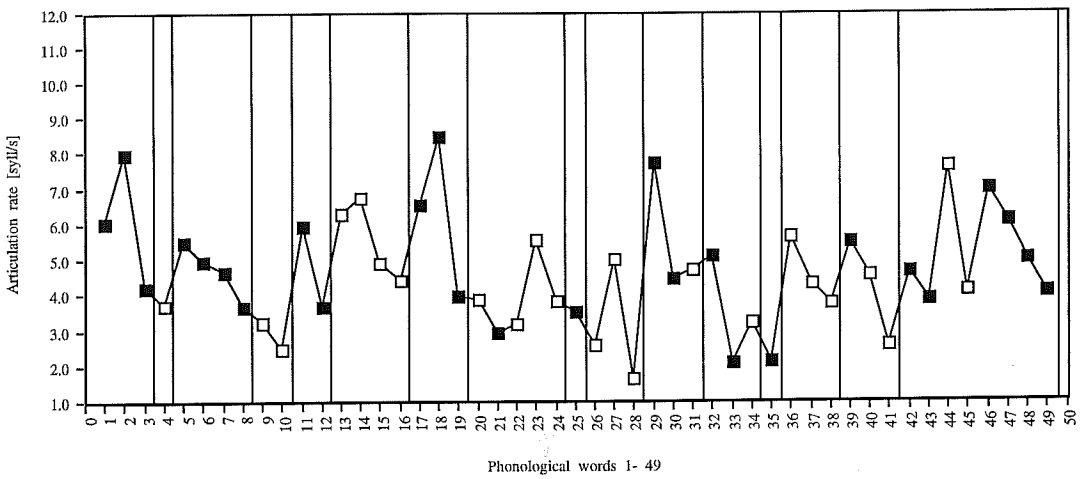
Four groups of interpause stretches were analysed: 4-word (7 interpause stretches within this group), 5-word (6), 6-word (5) and 8-word (6). The results of the GLM analysis were as follows. Factors of speaker, position of the word and its size in segments were non-significant for all the groups of interpause stretches. The same was true for size in syllables, apart from 8-word stretches, where this factor reached significance [F (5,48) = 3.86, p = 0.006]. The tendency was, of course, that the larger the number of syllables in the word, the faster its articulation rate. All interactions were non-significant. In general, these results suggested that interpause stretches do not involve any systematic pattern of articulation rate.

In order to give the reader a clearer idea of how articulation rate varies within interpause stretches, a set of graphs (Figure 3.7) is presented below, showing the flow of speech produced by one speaker, one phonological word after another (each word represented by a square), with their articulation rate on the y-axis and

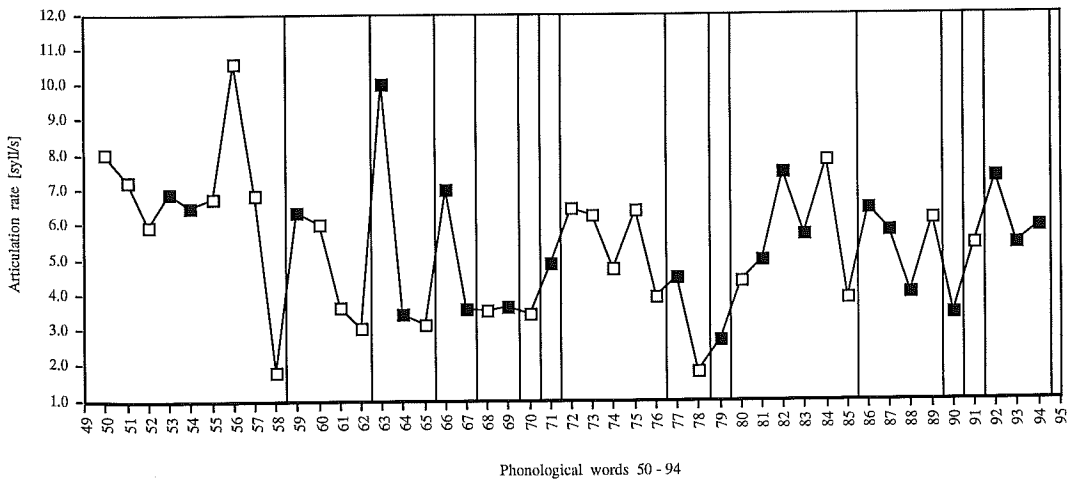
boundaries of interpause stretches marked by vertical lines. The graphs also include information about the boundaries of intonation phrases; the boundaries are indicated by the change of colour of the squares; successive squares of the same colour (white or black) belong to the same intonation phrase.

Figure 3.7 Articulation rate within interpause stretches (Speaker OL)

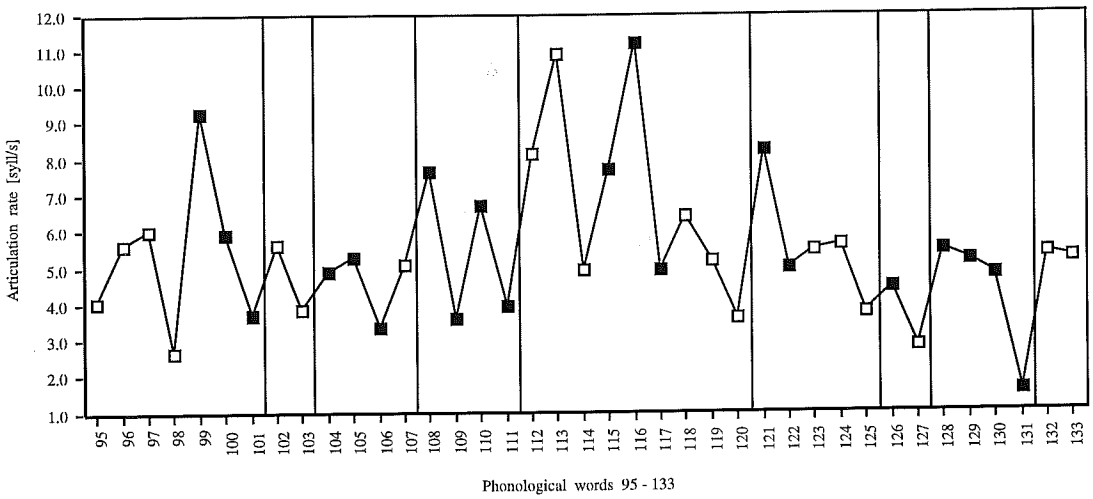
a) Interpause stretches containing phonological words 1- 49



b) Interpause stretches containing phonological words 50 – 94



c) Interpause stretches containing phonological words 95 - 133



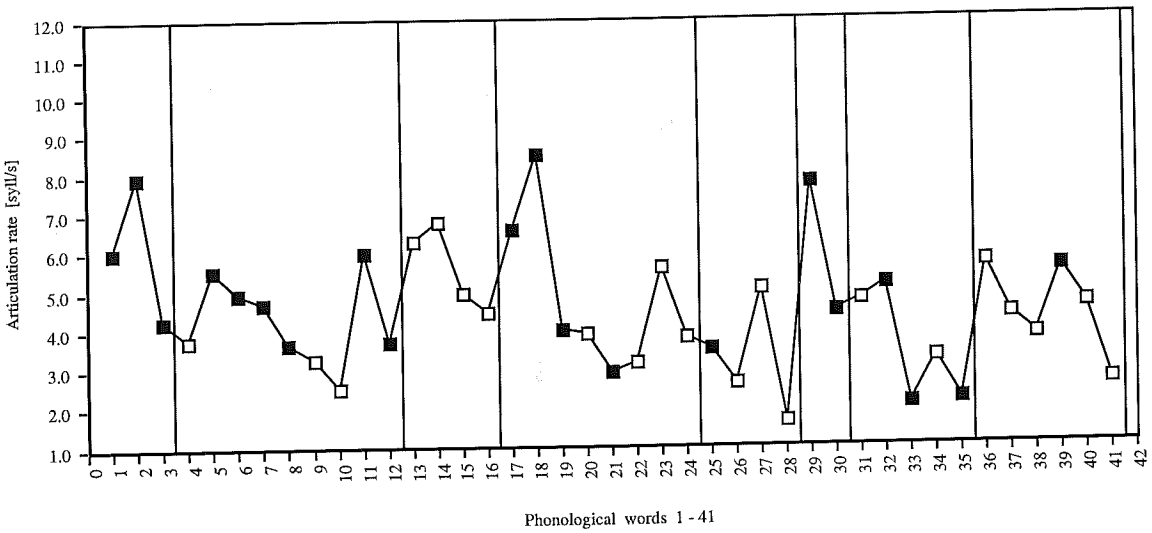
3.5.2 Clauses

Only three groups of clauses had a sufficient number of observations for the GLM analysis: 5-word (in total 7 clauses), 6-word (9) and 8-word clauses (7). The results were consistent across all these groups of clauses. The only significant effect was the factor of speaker in 8-word clauses [F (2,55) = 5.07, p = 0.01], but only two speakers, OL and JP, were responsible for this, having significantly different mean articulation rates. None of the interactions were significant. The absence of any significant results for position suggests that the clause is not a hierarchically higher domain of articulation rate variation either.

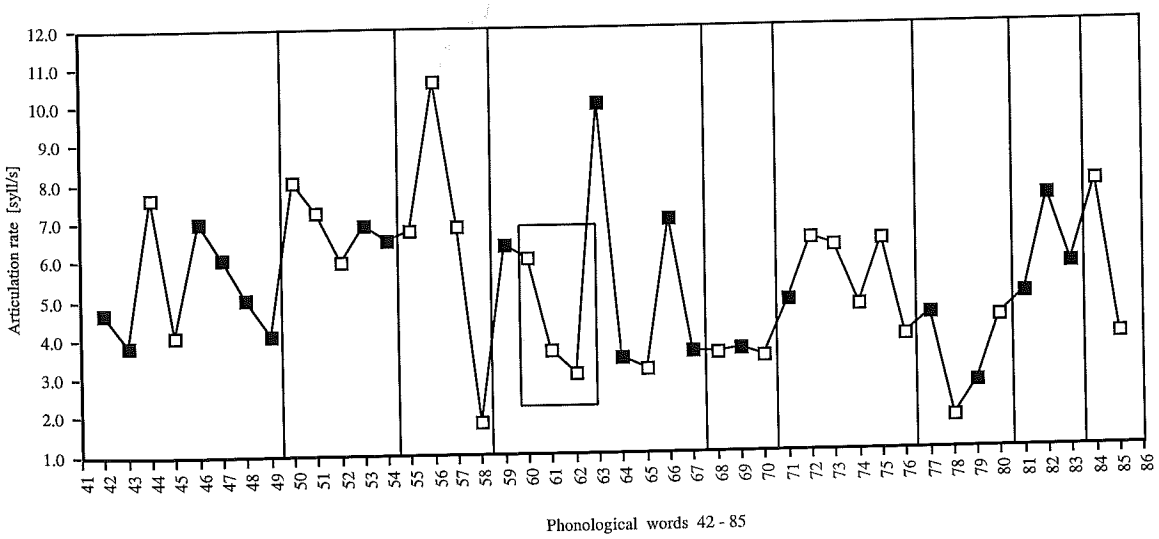
Figure 3.8 shows the variation in articulation rate in the speech sample from subject OL and Figure 3.9 in a part of the speech sample from subject JP. The principles of representation are exactly the same as for interpause stretches in Figure 3.7, with the exception of the vertical lines, which now indicate the clause boundaries. The frame around three words in Figure 3.8b represents an embedded relative clause.

Figure 3.8 **Articulation rate within clauses (Speaker OL)**

a) Clauses containing phonological words 1- 41



b) Clauses containing phonological words 42 - 85



c) Clauses containing phonological words 86 - 133

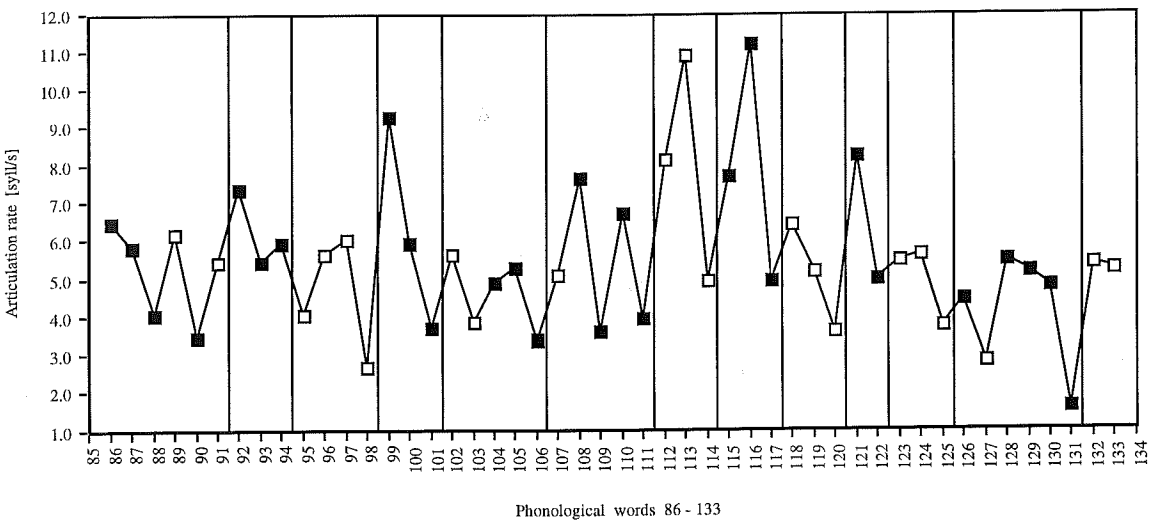
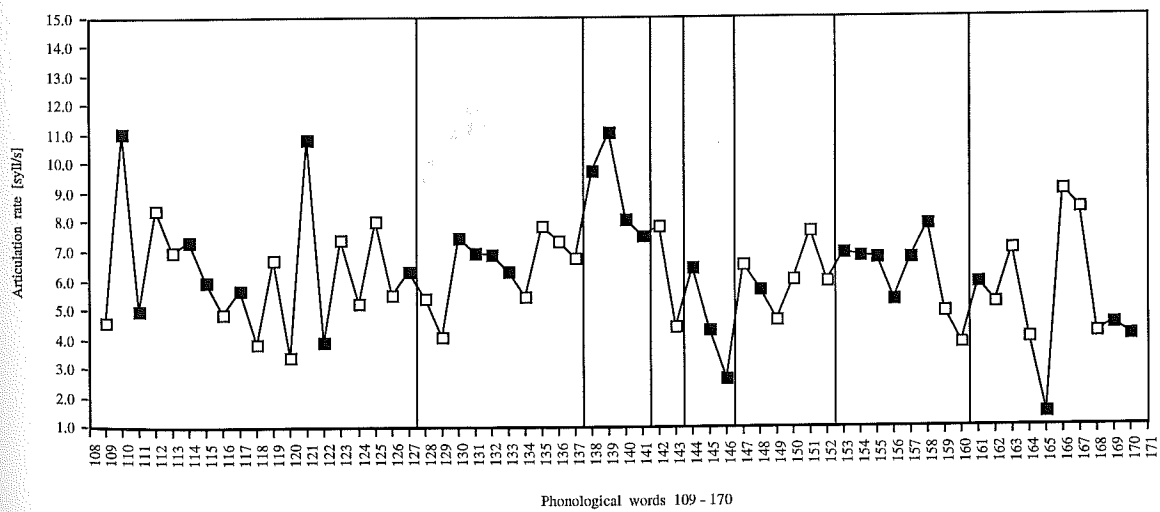


Figure 3.9 Articulation rate within clauses - words 109-170 (Speaker JP)



3.6 PHRASE-FINAL LENGTHENING

In the investigation of intonation phrases, final words stood out in the majority of cases as significantly different in their mean articulation rate from non-final words within the same phrase. As suggested above, this is likely to be an effect of phrase-final lengthening. The evidence presented here thus implies that the intonation phrase is also a domain of final lengthening. Since the presence of final lengthening can be considered a component of a systematic pattern of articulation rate variation, it was examined also in interpause stretches and clauses. It is possible that one or both of these units are higher-order domains of phrase-final lengthening. The hypothesis is that the last word within the interpause stretch or clause (which is also the last word in the final component phrase) is slower than the last word of the other phrases within the same unit. To test this, a paired-samples *t* test was carried out for the articulation rate of words at the end of non-final phrase(s) against the rate of words at the end of the final phrase.

3.6.1 Interpause stretches

The results of the analysis of phrase-final lengthening for interpause stretches are summarized in Table 3.10.

Table 3.10 ***t* values for tests of differences between articulation rates (AR) of final words in non-final intonation phrases within the interpause stretch, and of final words in the final intonation phrase within the interpause stretch (ns: $p > 0.01$).**

Speaker	AR [syll/s] Non-final IP	AR [syll/s] Final IP	*t* values	*df*	*p*	No. of interpause stretches
MS	5.86	4.22	-3.349	10	0.007	8
OL	4.43	3.67	-1.694	18	ns	14
JP	5.31	5.43	0.310	26	ns	16

Overall, the results were inconsistent. In two of the samples, phrase-final lengthening was greater in the last word of the final phrase than in the last words of non-final phrases within the interpause stretch, although significantly so only in the sample from MS. The sample from JP shows the opposite tendency but the trend is very small.

3.6.2 Clauses

The results presented in Table 3.11 do not indicate any systematic effect for phrase-final lengthening as a property of the clause. For two of the speakers, the articulation rate was, in fact, higher at the end the final phrase within the clause than at the end of clause-internal phrases, although significantly so only for one speaker.

Table 3.11 ***t* values for tests of differences between articulation rates (AR) of final words in non-final intonation phrases within the clause, and final words in the final intonation phrase within the clause (ns: $p > 0.01$).**

Speaker	AR [syll/s] Non-final IP	AR [syll/s] Final IP	*t* values	*df*	*p*	No. of clauses
MS	4.21	5.23	1.599	8	ns	7
OL	3.89	3.63	-0.945	25	ns	14
JP	4.85	5.79	2.773	33	0.009	13

3.7 CONCLUSIONS

In Study 1, six hypotheses were tested and the results can be summarized as follows.

Hypothesis 1.1 (null hypothesis):
Articulation rate is constant within the candidate domain.

As predicted, this hypothesis had to be rejected. Articulation rate was far from constant within the intonation phrase and, of course, also within interpause stretches and clauses since these either coincided with intonation phrases or contained more than one intonation phrase. The variation in articulation rate within the intonation phrase was typically between 20% and 40% from the mean.

For Hypothesis 2.1, a possibility of a recurring pattern in articulation rate was tested, as an alternative way of defining a domain of articulation rate variation.

Hypothesis 2.1:
Articulation rate variation within a candidate domain reflects a certain systematic pattern, which is the same in all three speakers.

This hypothesis was confirmed only for the intonation phrase. The pattern which seemed to dominate was rallentando, i.e., progressive slowing down. Neither interpause stretches nor clauses manifested any systematic pattern in rate.

Hypothesis 4.1 tested the overall statistical significance of this rallentando pattern and Hypothesis 5.1 tested its nature in more detail, namely, whether the progressive slowing down is linear or non-linear.

Hypothesis 4.1:

Articulation rate variation within the intonation phrase reflects a regular pattern of rallentando (progressive slowing down) in all three speakers.

Hypothesis 5.1:

The articulation rates of words in all positions within the intonation phrase are significantly different from each other in all three speakers.

Both Hypotheses 4.1 and 5.1 were confirmed for 2-word intonation phrases. For 3-word phrases, the rallentando pattern (ABC) was statistically significant only in one speaker, while for the other two speakers, the overall pattern turned out to be BAC. However, in none of the speakers were the first two positions statistically different. In the speaker with the rallentando pattern, this means that the progressive slowing down was non-linear. Only the final word was statistically different from the other words within the phrase. Thus Hypothesis 5.1 was not confirmed for 3-word phrases. Finally, for 4-word phrases, Hypothesis 4.1 was confirmed; the overall pattern was rallentando and there were no significant differences between speakers in this respect. However, the slowing down was non-linear; only the final word was significantly different from all the other words within the phrase; non-final words were not statistically different from each other. Thus Hypothesis 5.1 was not confirmed for 4-word intonation phrases either.

Hypothesis 3.1 concerned the question of determinants of articulation rate.

Hypothesis 3.1:

The size of the phonological word significantly affects the word's articulation rate in all three speakers.

It was confirmed for all three groups of intonation phrases but only when the word size was measured in syllables. The tendency was an increase in articulation rate

with an increasing number of syllables. When word size was measured in segments (with long vowels counted either as one or two segments), no apparent trend was found. Further investigation showed that this finding is due to a tendency of polysyllabic words in Czech to have a simpler syllable structure than monosyllables or disyllables.

The last hypothesis, Hypothesis 6.1, was not confirmed.

Hypothesis 6.1:
Speakers differ significantly from each other in their articulation rate.

In 2-word and 3-word phrases, the speakers were not significantly different from each other in their articulation rate. Only in 4-word phrases were two speakers statistically different.

CHAPTER 4

STUDY 2: EXPERIMENTAL DESIGN

4.1 INTRODUCTION

The results of Study 1 suggested that articulation rate in elicited spontaneous Czech speech does not vary in an arbitrary way but has a domain within which it varies systematically. This domain is the intonation phrase and the pattern of variation is a slowing down throughout the phrase. The analysis also showed that the slowing down is non-linear; it is rather gradual across non-final words, with the last word within the phrase being significantly slower than any other preceding words within the same phrase. The data set in this pilot experiment was too small to allow any strong conclusions about the consistency of the pattern across speakers and about the extent to which this pattern is typical for speech in general. Therefore, this pilot study was followed by Study 2.

The main aim of Study 2 was to test on a larger database, both in terms of the number of speakers (four this time) and the number of tasks (reading a story, paraphrasing a story, and spontaneous speech), the idea that the rallentando pattern within the intonation phrase is indeed the norm of articulation rate variation in Czech. The study thus concentrated entirely on the intonation phrase. Generally, I kept the structure of the analysis as similar as possible to that of Study 1, in order to make cross-experiment comparisons possible.

However, Study 2 also dealt with the issue of cross-speaker consistency in the articulation rate of individual phonological words. This issue was not investigated in Study 1 since the material used there was elicited spontaneous speech, i.e., samples with a different content from different subjects. The

material used in Study 2 included readings of the same story by different subjects and this made it possible to establish to what extent they agree in the variation of their articulation rate. This was tested as Hypothesis 1.2[1], by using a Pearson coefficient of correlation.

Hypothesis 1.2:
Speakers reading the same text tend to produce individual phonological words at similar rates.

The analysis in Study 1 based on ordinal patterns revealed some interesting facts about the relative distribution of all the patterns that occurred in the material; I used this analysis again in Study 2. However, the principal interest of this study was to test the statistical significance of the pattern of slowing down, in other words, to test the hypothesis introduced in section 3.4 as Hypothesis 4.1. Adjusted to the current material, this hypothesis is stated here as:

Hypothesis 2.2:
Articulation rate variation within the intonation phrase reflects a regular pattern of rallentando (progressive slowing down) in all four speakers and all three tasks.

Hypothesis 2.2 was tested by GLM analysis whose nature was described in more detail in section 2.7. Recall that from a statistical point of view, testing this hypothesis amounts to testing whether the factor of position of the word within the phrase has a significant effect on its articulation rate. The lack of a significant effect would indicate that the mean articulation rates of words within the intonation phrase are not statistically different from each other.

[1] Recall from p. 35 that the number 2 after the dot indicates that the hypothesis comes from Study 2.

As predicted, given the results of Study 1, a significant effect of word position did occur in the new data. Following the same procedure as in the previous study, the next step was testing the linearity of the progressive slowing down, by using a post-hoc Tukey test. In Study 1, the relevant hypothesis was Hypothesis 5.1, here adopted for the new material as:

Hypothesis 3.2:
The articulation rates of words in all positions within the intonation phrase are significantly different from each other in all four speakers and all three tasks.

In Study 1, the test of possible factors influencing the pattern of articulation rate variation suggested that the size of the phonological word, expressed in the number syllables, significantly affects how fast the word is pronounced. In Study 2, a similar analysis was performed to find out whether the same holds for the new material. The hypothesis testing word size is Hypothesis 4.2.

Hypothesis 4.2:
The size of the word, counted in syllables, significantly affects the word's articulation rate in all four speakers and all three tasks.

As discussed in the literature survey, in section 1.3.5, one of the factors which seem to affect the articulation rate of the word is whether it belongs to the category of content or function words, function words being generally faster than content words (Kaiki, Takeda & Sagisaka, 1990; Fant, Kruckenberg & Nord, 1991a; O'Shaughnessy, 1995). In Study 2, I included the factor of word class into the analysis as well, in order to test whether the same observation applies to Czech. The hypothesis concerning word class was:

Hypothesis 5.2:

The word class significantly affects the word's articulation rate in all four speakers and all three tasks.

In Study 1, the speakers did not seem to differ significantly from each other in their mean articulation rates. Since this might be a consequence of analysing a relatively small number of observations, Hypothesis 6.1 from the previous study was adopted and tested for the new material as:

Hypothesis 6.2:

Speakers differ significantly from each other in their articulation rate of phonological words in all three tasks.

As for the task factor, a faster articulation rate has often been demonstrated in different languages for reading as opposed to spontaneous speech (Lehtonen, 1979; Butcher, 1981; Strangert, 1993). Finding the same results for Czech would thus not be a great surprise. The material used in Study 2 includes also paraphrasing the story. As will be described in section 4.2, paraphrasing shares a number of cognitive activities with spontaneous speech. However, it is likely that the retrieval of the story's content from memory makes the task more difficult and thus people might speak more slowly than they do in their spontaneous speech. The hypothesis about task is as follows:

Hypothesis 7.2:

The words' articulation rate is significantly affected by the kind of task the speaker is carrying out. Reading is significantly faster than spontaneous speech, and spontaneous speech is faster than paraphrasing.

Hypotheses 2.2-7.2 were tested simultaneously in a GLM analysis, by including position, word size, word class, speaker and task as explanatory variables, and by using a post-hoc Tukey test.

4.2 MATERIAL

The main intention in the choice of material and subjects was to obtain data as balanced as possible. With respect to material, three tasks were selected: reading a story, paraphrasing the story and elicited spontaneous speech. These tasks differ in terms of what cognitive activities are involved in them (Slis *et al.*, 1979; Levelt, 1989). Reading does not involve message or content generation, creative grammatical generation, and the selection of words. The main demands are understanding and interpreting the text and translating the orthographic form into an articulatory plan for the spoken form. Paraphrasing the story is a more complex task. Although the content or story is given, the speaker has to retrieve it from his memory, form his own sentence structure and select words before he is ready to produce a spoken utterance. Producing spontaneous speech shares most of the planning stages with paraphrasing, but it also includes an active decision about the content and considering the pragmatic aspects of the utterance.

For the reading task, a short story from a book by a well-known Czech writer Ota Pavel was chosen.[2] The story is written in a relaxed, semi-colloquial style and so it was easy to read. The whole text took on average 5-6 minutes to read. However, only the middle third of the story, containing 295 words, was analysed in this study (the text, together with a translation into English, is presented in Appendix 5). The first third was not chosen since some speakers seemed still somewhat nervous and made an impression of speaking at a slower rate than was usual for them. The final third of the story was left out from the material because a number

[2] Ota Pavel (1974) *Jak jsem potkal ryby*. Prague: Mladá fronta.

of subjects seemed to be getting tired and so the number of slips of the tongue rose.

Paraphrasing of the story was included in the material in its full length for each speaker. Regarding elicited spontaneous speech, I recorded between 7 and 14 minutes of speech per subject. My preference was to use parts of the speech which (i) were not interrupted by my questions (see the discussion in section 1.3.4), and (ii) which seemed most natural in the sense of matching closely the subject's speech outside the recording situation. These criteria compelled me to cut down the length of the spontaneous speech samples to a maximum of about 4 minutes per speaker.

The entire material consists of 33 minutes and 16 seconds of speech, counting all speakers and tasks together. Details about the precise duration of the speech samples used for analysis are given in Table 4.1 below. In all but one sample of spontaneous speech, the time given is the time during which a subject talked without any interruption. The sample of spontaneous speech from the subject PK did contain an interruption; the duration of his speech before and after the interruption is given in brackets.

Table 4.1 The duration of speech samples used in the study.

Speaker	Reading [min]	Retelling [min]	Spontaneous speech [min]	Total [min]
AV	1:58	2:11	2:53	7:02
LG	2:05	1:49	3:51	8:45
OU	1:42	3:39	3:38	8:59
PK	1:46	3:57	2:47 (1:18 + 1:29)	8:30
Total time	7:31	11:36	13:09	33:16

4.3 SUBJECTS

A total number of ten speakers took part in the recording, although only four of them were chosen for the present study; see below in this section. The criterion for the selection of all subjects was to assemble a group as homogeneous as possible as to age, sex, education, dialect, social background, speech task, and familiarity with the interlocutor, in this case, myself. A similar criterion was used in Study 1; in the description of that experiment in Chapter 2, I stated reasons behind this criterion.

All subjects were native speakers of Czech, without a strong regional dialect, in the age category between 23 and 34 years and with a university education. They came from families which could be characterized as middle class in the British sense. The degree of familiarity with the interlocutor was also uniform; most of the subjects were my friends. This made it easier to obtain a natural performance in the elicited spontaneous task than if I had been a complete stranger to the speakers. None of the speakers reported any speech or hearing problems. They did not know the exact purposes of the experiment.

As mentioned above, only four speakers were in the end selected for the study, two male (LG and PK) and two female (AV and OU). The principal reason for reducing the number of subjects was that the database would be too large, given the scope of the study and the available computing capacity, if all ten speakers were to be included in the analysis. The main criteria for excluding particular subjects were a larger than average number of slips of the tongue in reading, a reduced ability to relax in a given situation and produce spontaneous speech (one speaker, for instance, would only speak in a lecture-like manner in front of the microphone), and/or the fact that some speakers were simply not very talkative. The first two, in particular, may be due to nervousness. To exclude speakers who seemed rather nervous is, I believe, justifiable for the present study: it is possible that nervousness affects articulation rate and also frequency and duration of

pauses, although it is not clear in what way. The four speakers selected for the analysis seemed comparably relaxed and produced speech which sounded like their normal way of speaking.

4.4 RECORDING PROCEDURE

The recording took place in the recording studio of the Phonetics Institute, Charles University, Prague, with a high quality Marantz professional cassette recorder and a Pioneer DM-21 microphone. Every subject was recorded solely in my presence.

The subjects were first asked to read twice a set of twenty sentences intended for another experiment. This was followed by the reading task, at the beginning of which everybody had a short time to read the story for himself/herself. The subjects then proceeded to paraphrasing the story, for which a short time for preparation was also allowed. The instruction was to reproduce the story in a reasonable detail (i.e., not just briefly summarizing it), in a form that the speaker was comfortable with.

Finally, elicited spontaneous speech was recorded. By this time, most of the subjects seemed relaxed enough to give a close approximation of their normal conversational speech. The aim was to make the subject speak freely for as long a time as possible without any interruptions, but without explicitly asking for a monologue. The reason was to avoid as much as possible additional, discourse factors of a dialogue, which could affect the temporal structure of the subject's speech. To prompt this, the initial question was typically about the subjects' everyday life, their children, home, holidays, etc.

The total time of the recording for a single subject did not exceed 30 minutes.

4.5 METHODOLOGY FOR THE ANALYSIS OF ARTICULATION RATE VARIATION WITHIN THE INTONATION PHRASE

Since one of the main aims of this study was to test the results of Study 1 on more extensive data that would at the same time include materials of a different kind, the methodology was essentially the same as in the first study. I shall, therefore, confine myself only to a necessary description and some additional remarks.

Recall from section 2.5.1.3 that the definition of the intonation phrase accepted in this thesis states that the intonation phrase is a stretch of speech determined by the intonation contour. As in Study 1, only one level of intonational phrasing was used in Study 2.

The phonological word was employed again as the unit for the measurement of articulation rate within the intonation phrase. The reasons were explained in section 2.5.3. An additional justification for using the phonological word in Study 2 was that it proved to be a good choice in Study 1; choosing the phonological word as the minimum unit for measurement led to discovering a systematic pattern in articulation rate variation.

Articulation rate was again expressed in terms of number of syllables per second. Unlike in Study 1, where all three speakers were rather precise in their pronunciation, in this study a decision had to be taken as to whether to calculate the number of syllables in the word according to its 'surface form' or 'underlying form'. My decision was to count the number of syllables according to the 'surface form', where this form was fully lexicalized in colloquial Czech, and according to the 'underlying form' in other cases. The group of fully lexicalized items involved, more or less exclusively, words such as *'ňák'* ('in some way') *'ňáký'* ('some') and its inflexional relatives.

A special treatment for non-lexicalized items (or rather - item - since only one occurred; the trisyllabic *'protože'* - 'because' - was pronounced as disyllabic *'-tože'*) was chosen since the syllable deletion in this case is likely to be a result of a faster articulation rate. By not counting the unpronounced syllable, one would ignore one of the manifestations of an increased articulation rate.

4.6 MEASUREMENT TECHNIQUE

Speech samples were digitised at a resolution of 16 bits and a sampling rate of 11 kHz using the Entropics Waves+ speech analysis software on a Sun Sparcstation LX. In order to make measurements of the duration of phonological words as accurate as possible, a combination of impressionistic listening and reference to waveform displays and wideband spectrograms was used.

Both phonological words and intonation phrases were identified impressionistically by myself. In the case of intonation phrases, an independent analysis of the text was obtained also from another Czech phonetician to avoid idiosyncrasy in determining phrase boundaries. As in Study 1, the agreement between the two analyses was very high; the few cases of disagreement were again adjudicated by a third phonetician.

Phonological words were identified only by myself. There were only a few cases where I was in doubt about whether the function word was an enclitic or a separate phonological word; another phonetician was consulted about these.

4.7 DETERMINANTS OF ARTICULATION RATE VARIATION

In the first chapter, section 1.3.5, I gave a survey of the factors reported in the literature as significant for variation in articulation rate. They include phonological and syntactic characteristics of the unit for which articulation rate was measured and its function in discourse. Phonological characteristics, represented by word size, were tested already in Study 1, together with the factor of the word's position within the intonation phrase, which was included to test the statistical significance of a progressive slowing down. As I mentioned in section 4.1, in Study 2, the factor of position and the factor of word size were included in the analysis again, in order to test the findings from Study 1. Apart from these two factors, the content/function word dichotomy was considered too, to represent syntactic characteristics of the words and to test the findings reported in the literature for other languages.

4.7.1 Position of the word within the intonation phrase

Words were coded as being in the first, second, third, or fourth positions within the intonation phrase, and a statistical analysis was carried out to find out whether the word's position is a factor significantly affecting the pattern of articulation rate variation.

4.7.2 Size of the phonological word

The size of the word was expressed only in terms of the number of syllables in Study 2, since when the number of segments was used in Study 1, no significant effect of word size on the word's articulation rate was found.

4.7.3 Function vs. content words

The studies reporting an effect of content/function word dichotomy very rarely state explicitly what they consider a function or a content word (e.g., Fant, Kruckenberg & Nord, 1991a; O'Shaughnessy, 1995). In fact, when one goes beyond basic textbooks on morphology, syntax and semantics, classifying words into content and function word categories is not as straightforward as we would wish.

For the purpose of the present study, a traditional definition of function and content words was used (Quirk & Greenbaum, 1973; Bolinger, 1975; Crystal, 1980; Radford, 1988; Finegan, 1994):

- content (lexical) words - an open class of words with stateable lexical meaning
- function (grammatical) words - a closed class of words used to express grammatical relationships

The group of content words includes nouns, adjectives, verbs, adverbs and also numerals; the group of function words includes pronouns and generally pro-forms, auxiliary verbs, prepositions and conjunctions. Articles also belong to the function word category but they are irrelevant for this study since they do not occur in Czech. My discussion with a number of syntacticians revealed that there is disagreement particularly about how to classify pronouns, auxiliary verbs and also prepositions. The decision depends largely on the syntactic theory.

The division of word classes into the two categories as outlined above seems to be the most traditional one and will be used in the present study. However, the unit for measuring articulation rate is the phonological word and the correspondence between phonological words and content and, especially, function words is not straightforward.

With respect to prosodic structure, content words are normally stressed and thus they form phonological words. Depending on discourse structure, tempo, etc., some content words, especially monosyllabic ones, may cliticize, i.e., form a single phonological word with a neighbouring word, usually the preceding one. These cases are relatively rare – a single content word per phonological word is the most common 'type' of the phonological word.

As to function words, several possibilities arise, depending on the discourse structure. Function words may be prosodically independent - they may have stress and form a phonological word on their own. However, often they lose their stress in continuous speech, particularly the monosyllabic ones (Roach, 1983; Palková, 1994), and cliticize with their neighbouring words. Normally they cliticize with the preceding ones in which case the function word becomes an enclitic. A monosyllabic function word may also cliticize with the following word, i.e., become a proclitic. This might happen particularly at the beginning of an intonation phrase. In some instances, several function words follow each other, in which case the first one tends to be stressed and the following ones are enclitics. Together they form a phonological word. Some examples of content vs. function words in relation to the placement of stress were presented in section 2.5.4.

In my investigation, I aimed for such a classification of phonological words in relation to the content vs. function word dichotomy, which (i) would preserve this two-way distinction as much as possible, and, at the same time, (ii) would not involve too many 'types' of phonological words, since this would lead to having too small a number of observations within some groups. I thus designed a three-way classification of phonological words:

(1) phonological words containing one or more content words (no function words)

(2) phonological words containing a content word and one or more (cliticizing) function words

(3) phonological words containing one or more function words (no content word)

The factor which this classification represents will be referred to as 'word class'.

4.8 STATISTICAL PROCESSING

The statistical processing was essentially the same as in Study 1. The GLM procedure (General Factorial) was employed, with articulation rate being the dependent variable, and position, word size, word class, speaker and task being the explanatory variables. When a factor achieved an overall statistical significance, the GLM analysis was combined with a post-hoc Tukey test, in order to find out which levels or categories within the factor were statistically different from each other.

All the explanatory variables were used as factors, although, technically speaking, only word class, speaker and task were nominal variables, while position and word size were numerical variables. However, as explained in section 2.7, coding position and word size as covariates, rather than factors, would prevent me from conducting a post-hoc Tukey test, crucial for the analysis.

The statistical analysis also included the coefficient of determination (R^2), which provided an estimate of how much variance in the dependent variable can be explained by the model in terms of the factors incorporated in the model. The analysis was executed again separately on groups of 2-word, 3-word and 4-word intonation phrases. In order to be able to compare the success of the model across

these three groups of phrases, which involved a different number of parameters (factors), the adjusted R^2 was used as well. The adjusted R^2 has the advantage of giving an estimate while weighting according to the number of parameters in the model.

In Study 2, I also addressed the issue of which factors affected the articulation rate of phonological words to the greatest extent. To estimate relative importance of the individual factors, a measure of the strength of association, η^2, was calculated. The formula for calculation is:

$$\eta^2 = SS_x / SS_{Total}$$

where SS_x is the sum of squares of any factor effect or the interaction of any factor effects, and SS_{Total} represents the total sum of squares (Linton & Gallo, 1975).

The significance level used in Study 2 is the same as in Study 1, i.e., $p < 0.05$.

All the statistical processing was carried out using the same statistical package as in Study 1, i.e., SPSS for Windows, version 7.5.1.

CHAPTER 5

STUDY 2: ELEMENTARY RESULTS

5.1 FREQUENCY DISTRIBUTION OF INTONATION PHRASES OF DIFFERENT SIZES

As described in section 4.6, in each speech sample the boundaries of phonological words and intonation phrases were identified impressionistically. Table 5.1 below shows how many intonation phrases fell into each category, the category being determined by the number of phonological words the phrase contained. Since the number of phrases consisting of five, six and seven words was generally too small to include these phrases into the analysis, these categories were collapsed into one (5+) in the table below. The columns with values in bold (2-word, 3-word and 4-word phrases) represent the categories of phrases on which Study 2 was based. Intonation phrases which contained a slip of the tongue, misreading or a correction were excluded from the analyses and also from Table 5.1.

The total number of phonological words measured was 2,858 (reading: 714; paraphrasing: 909; spontaneous speech: 1,235). However, in the main analysis of Study 2, only 2-word, 3-word and 4-word intonation phrases were examined. The total number of words used in this main analysis was 2,102; the total number of phrases investigated was 784 (209 in the reading task, 255 in the paraphrasing task, and 321 in spontaneous speech task.).

Table 5.1 **The frequency distribution of intonation phrases in categories of 1-word, 2-word, 3-word, 4-word and 5-or-more-word intonation phrases. (The categories in bold represent intonation phrases which were used for the analyses in the present study.)**

READING

Speaker	Categories of IP's (Number of words in the IP)					Total no. IP's
	1	**2**	**3**	**4**	5+	
AV	8	**24**	**23**	**7**	5	67
LG	8	**21**	**24**	**7**	6	66
OU	7	**23**	**22**	**7**	6	65
PK	7	**23**	**22**	**6**	6	64
Total	30	**91**	**91**	**27**	23	262

PARAPHRASING

Categories of IP's (Number of words in the IP)					Total no. IP's
1	**2**	**3**	**4**	5+	
17	**30**	**14**	**7**	3	71
41	**25**	**16**	**6**	1	89
67	**34**	**28**	**10**	4	143
71	**44**	**36**	**5**	2	158
196	**133**	**94**	**28**	10	461

SPONTANEOUS SPEECH

Speaker	Categories of IP's (Number of words in the IP)					Total no. IP's
	1	**2**	**3**	**4**	5+	
AV	27	**31**	**26**	**12**	9	105
LG	54	**40**	**39**	**14**	7	154
OU	56	**32**	**29**	**17**	8	142
PK	52	**38**	**29**	**14**	8	141
Total	189	**141**	**123**	**57**	32	542

Table 5.1 also shows a difference between the tasks in the frequency distribution of phrases according to their size. In reading, the most common phrases were those containing two or three words; in paraphrasing and spontaneous speech, the prevailing number of phrases contained only one word. This suggests, not surprisingly, that one of the important factors for the speaker in placing intonation phrase boundaries is hesitation. Phrases with two or three words were generally very common in paraphrasing and spontaneous speech too. Phrases containing more than three words were progressively less common, with 6-word and 7-word phrases being exceptional. Comparison with other data on Czech (e.g., Palková, 1974) revealed that this distribution is entirely typical.

5.2 INTER-SPEAKER CORRELATION OF ARTICULATION RATE OF PHONOLOGICAL WORDS (READING)

As mentioned in section 4.1, having available recordings of the same text read by different speakers made it possible to carry out cross-speaker comparisons of the articulation rate of individual phonological words. The hypothesis was formulated as follows:

> **Hypothesis 1.2:**
> **Speakers reading the same text tend to produce individual phonological words at similar rates.**

This hypothesis incorporates the assumption that, in reading, speakers tend to place phonological word boundaries in the same positions. The analysis of the material confirmed this assumption. Only in two cases were the phonological word boundaries placed differently by one of the speakers. In both cases, the whole intonation phrase containing the word in question was excluded from the analysis. The intonation phrases in which at least one of the speakers made a slip of the tongue or misread the text were also excluded (there were 10 of these altogether). A total number of 54 intonation phrases were included in the analysis. In terms of the number of words, 81% of the 193 phonological words in the passage were used in the analysis.

Prior to the statistical analysis, a graphical display (Figure 5.1) was used to observe how closely the speakers agree in the variation in their articulation rate when reading the same text. Phonological words are placed on the x-axis in the order in which they occurred in the speech samples and they are presented in Czech orthography; the articulation rate is on the y-axis. The descending black lines indicate where in the text the excluded intonation phrases belong. Intonation phrase boundaries are marked by broken vertical blue lines.

Figure 5.1 Articulation rate variation in reading: individual speakers.

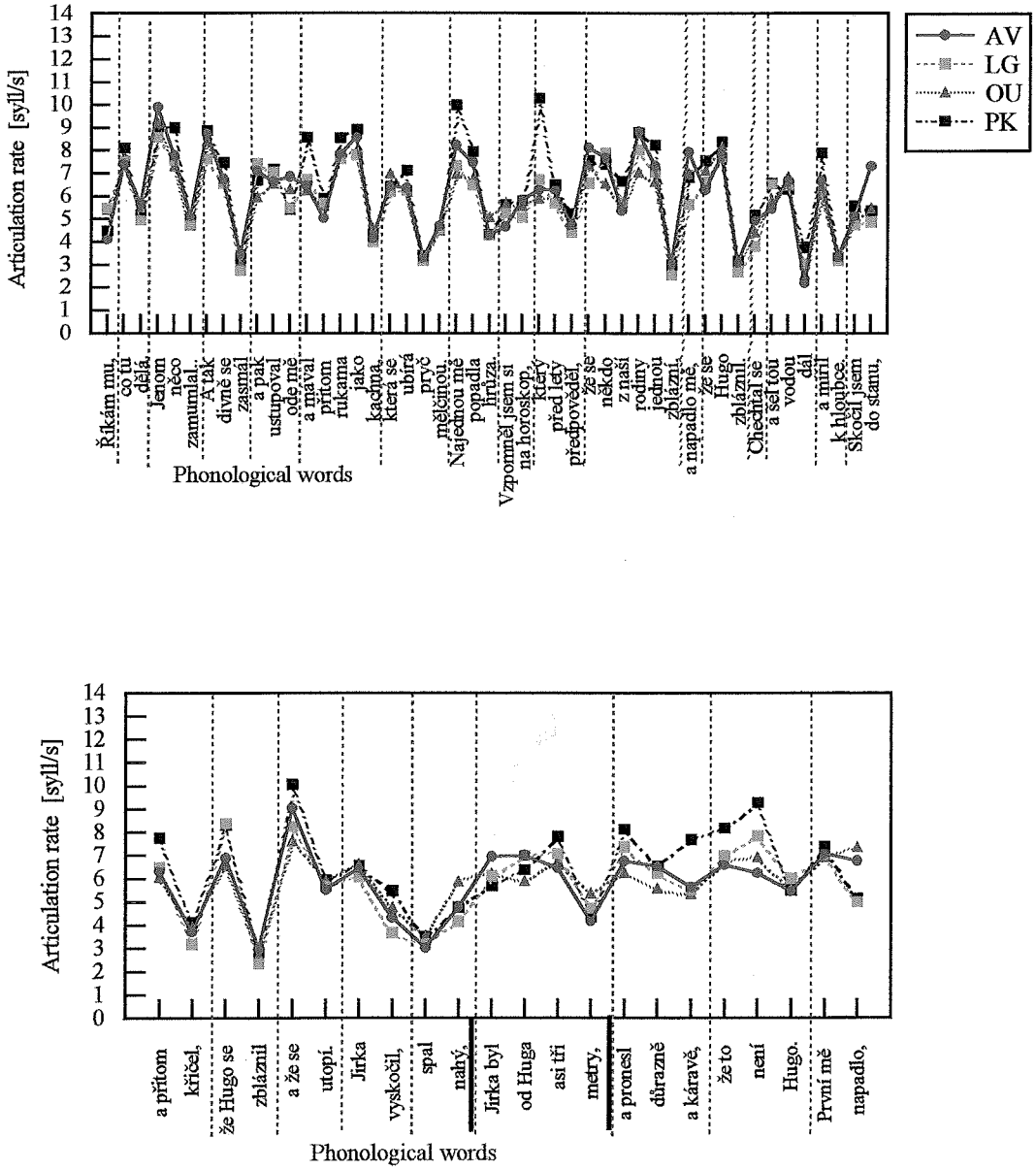


Figure 5.1 Articulation rate variation in reading: individual speakers. (cont.)

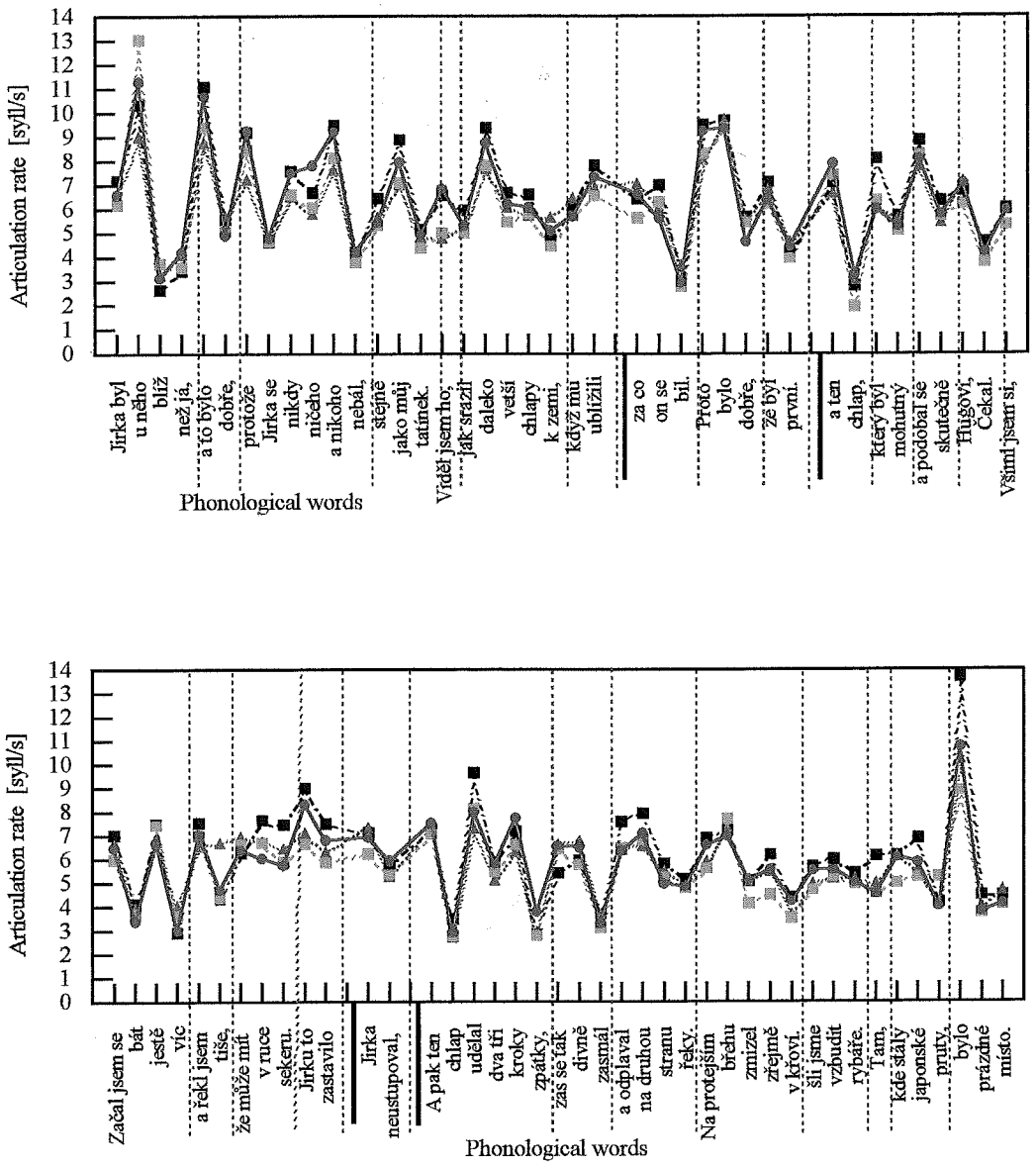


Figure 5.1 illustrates that speakers do tend to produce the same phonological words at a similar articulation rate. The agreement is rather striking. A statistical analysis, using Pearson coefficient of correlation *r*, was carried out to quantify the strength of the relationship between individual speakers' articulation rates. Table 5.2 below shows the results in the form of a matrix of correlations between the four speakers.

Table 5.2 **Inter-speaker correlations of the articulation rate of phonological words in reading (Pearson's correlation coefficient *r*).**

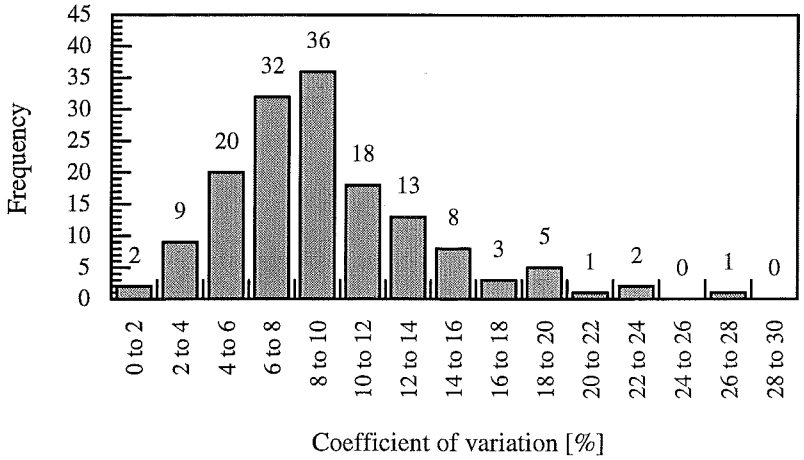
Speakers	AV	LG	OU
LG	0.915		
OU	0.917	0.903	
PK	0.902	0.905	0.876

Number of observations: 150 for each speaker

All coefficients of correlations in Table 5.2 are significant at 0.1% level. Hypothesis 1.2, that speakers reading the same text tend to produce individual phonological words at similar rates, was confirmed.

A further analysis, using the coefficient of variation (percentage of deviation from the mean), was conducted to quantify the dispersion of articulation rates for each individual phonological word as produced by the four speakers. Figure 5.2 presents a frequency distribution of phonological words according to their coefficient of variation.

Figure 5.2 **Frequency distribution of phonological words according to their degree of variation across different speakers.**



Oxford University
Phonetics Laboratory

Out of 150 phonological words, only 15 words had a coefficient of variation greater than 15% and in none of the words did the variation exceed 28%. In the majority of words (66% of the total), the coefficient of variation was below 10%. Evaluating the degree of dispersion based on the coefficient of variation is not straightforward because there does not seem to be an established criterion for what is regarded as a significant variation in articulation rate. However, a threshold around 20% can be discerned from the conclusions (unfortunately, typically lacking in theoretical motivation or discussion) of previous studies (Goldman-Eisler, 1968; Grosjean & Deschamps, 1975; Miller *et al.*, 1984). Variation below 10% tends to be regarded as non-significant.

All the phonological words for which the variation across speakers exceeded 15% were examined more closely in order to find out what common properties, if any, they have. I focused on the words' position within the intonation phrase and the word class as defined in section 4.7.3. The 15 words belong to different intonation phrases.

The examination showed a relatively consistent trend as to the position of the word but none as to the word class. Regarding the factor of position, 9 out of 15 words were in the last position in the intonation phrase, 2 in an 'internal' position, and 4 in the first position. As to word class, 6 phonological words were single content words, 5 phonological words fell into the category of a content word plus one or more function words, and 4 phonological words were function words.

Summary:
When speakers read the same text, they tend to produce individual phonological words at similar rates. The variation across speakers rarely exceeds 15% from the mean and the words concerned tend to be in the last position within the intonation phrase. A possible, although speculative, explanation might be that speakers use phrase-final lengthening to a different degree.

The high correlation between different speakers' articulation rates in reading indicates strongly that the articulation rate of phonological words is not arbitrary. It follows that the patterns of rate variation throughout the speech of individuals cannot be arbitrary either. There must be some underlying factors responsible for this. Some plausible candidates are explored in Chapter 6.

5.3 PATTERNING OF ARTICULATION RATE VARIATION WITHIN THE INTONATION PHRASE

As I claimed in section 3.3, the analysis of ordinal patterns in Study 1 revealed a number of interesting facts about the relative distribution of all the patterns that occurred in the material. I shall use the same kind of analysis for the present material, to be able to establish how consistent the patterning of articulation rate variation is in speech.

The results of the analysis of ordinal patterns on 2-word, 3-word and 4-word intonation phrases are presented in Tables 5.3 (2-word phrases), 5.4 (3-word phrases) and 5.5 (4-word phrases). The representation works on the same principles as in the first study (see section 3.3).

As mentioned in section 5.1, the intonation phrases which contained a slip of the tongue, stutter, or, in reading, a misreading, were not used in this, and the following, analyses. For the analysis of inter-speaker correlation of articulation rate in reading, described in section 5.2, the same number of observations was necessary across the speakers. Thus the intonation phrases in which even only one subject misread the text were excluded from all the subjects' samples. This is not necessary in the present analysis and the rest of the study, and thus the number of intonation phrases in different speakers may be unequal.

5.3.1 Two-word intonation phrases

Table 5.3 below shows, for each of the three tasks, the distribution of 2-word phrases according to their pattern of articulation rate variation.

In all three tasks there was a very strong tendency for the second word to be pronounced more slowly than the first one (83% of the total number of phrases in reading, 89% in paraphrasing and 79% in spontaneous speech).

Table 5.3 **Ordinal patterns for 2-word intonation phrases. (R = reading, P = paraphrasing, S = spontaneous speech)**

	AB 			BA 		
Speaker	R	P	S	R	P	S
AV	20	29	26	4	1	5
LG	18	19	31	3	6	9
OU	18	31	26	5	3	6
PK	20	39	28	3	5	10
Total	76	118	111	15	15	30

5.3.2 Three-word intonation phrases

Table 5.4 demonstrates that the three tasks were again very similar as to the kind of patterns which occurred and the number of intonation phrases representing them.

In 3-word phrases the fastest word was the first or the second one (reading: 98%, paraphrasing: 98%, spontaneous speech: 96%); the highest articulation rate in the last word was exceptional. In fact, the last word was usually the slowest (reading: 82%, paraphrasing: 78%, spontaneous speech: 73%).

Table 5.4 **Ordinal patterns for 3-word intonation phrases. (R = reading, P = paraphrasing, S = spontaneous speech)**

	ABC 			BAC 			BCA 		
Speaker	R	P	S	R	P	S	R	P	S
AV	12	7	10	8	5	11	-	2	1
LG	10	2	13	8	8	12	2	2	5
OU	10	12	15	7	10	9	1	3	1
PK	7	17	14	13	12	6	-	3	2
Total	39	38	52	36	35	38	3	10	9

	ACB 			CAB 			CBA 		
Speaker	R	P	S	R	P	S	R	P	S
AV	3	-	4	-	-	-	-	-	-
LG	3	3	8	-	-	1	1	1	-
OU	3	2	1	-	1	2	1	-	1
PK	2	4	6	-	-	-	-	-	1
Total	11	9	19	-	1	3	2	1	2

Out of six possible patterns, five occurred in reading and all six were found in paraphrasing and spontaneous speech. The distribution of intonation phrases within these patterns was uneven although in a highly similar way for all three tasks. The prevailing patterns were always ABC, i.e., decreasing articulation rate throughout the phrase (43% of the total number of phrases in reading, 40% in paraphrasing, and 42% in spontaneous speech), and BAC, where there is an increase in rate first and then decrease towards the end of the phrase (40% in reading, 37% in paraphrasing, and 31% spontaneous speech). Pattern ABC seems to be the strongest in spontaneous speech but this is due only to speakers OU and PK; in the other two speech samples the number of phrases with patterns ABC and BAC is balanced. In paraphrasing and reading, the overall tendency for pattern ABC to be slightly stronger than pattern BAC reflects the situation for all

but two speakers, for whom pattern BAC was the most typical one (speaker LG in paraphrasing, and speaker PK in reading).

5.3.3 Four-word intonation phrases

In 4-word phrases (Table 5.5), the trend in all three tasks was similar to that in 3-word phrases. The first or the second word tended to be the fastest in the phrase (59% of the total number of phrases in reading, 75% in paraphrasing, and 74% in spontaneous speech). In some cases, the third word was the fastest (reading: 33%, paraphrasing: 21%, spontaneous speech: 26%). It was extremely rare for the last word within the 4-word phrase to be pronounced with the highest articulation rate. There were no such cases in reading and spontaneous speech and only one case occurred in paraphrasing. Again, the last word tended to be the slowest (reading: 70%, paraphrasing: 79%, spontaneous speech: 68%).

As far as the overall structure of articulation rate is concerned, four words within the phrase give 24 possible patterns. However, not all of them occurred in the material. When the tasks were considered as one group, the number of patterns altogether was 19, with the 'missing' patterns being those in which the last word in the phrase (D) would have the highest articulation rate. This is not surprising since, as noted above, the last word tends to be the slowest within the phrase.

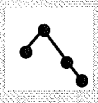
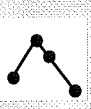
Considering the tasks separately, there were 10 patterns in reading, 11 in paraphrasing and 16 in spontaneous speech. The larger number of patterns in spontaneous speech is likely to reflect the double number of 4-word phrases in this task as opposed to reading and paraphrasing.

The number of phrases representing the individual patterns was uneven; the most common patterns (when not differentiating the tasks) were ABCD, BACD, BCAD, ACBD, CABD and CBAD, with at least 10 phrases per pattern. These six

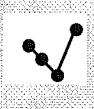
patterns are, in fact, identical to the 3-word patterns, with the addition of one more word, which is always the slowest of all the words within the phrase. The remaining patterns were represented very weakly, often only by a single phrase.

The analysis of individual tasks with respect to the distribution of phrases in the patterns cannot lead to any strong conclusions since the number of tokens within the categories is too small. Perhaps just one remark holds: with the exception of patterns ABCD, which is very weak in reading, and BCAD, which was not represented in reading at all, the overall observation about the typicality of the six patterns named above applies to all three tasks.

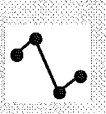
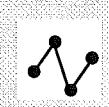
Table 5.5 **Ordinal patterns for 4-word intonation phrases. (R = reading, P = paraphrasing, S = spontaneous speech)**

	ABCD rallentando 			**BACD rallentando** 			**BCAD rallentando** 			**ACBD rallentando** 		
Speaker	R	P	S	R	P	S	R	P	S	R	P	S
AV	1	-	-	3	2	-	-	-	2	1	4	1
LG	-	-	4	1	1	2	-	1	-	1	-	2
OU	-	2	2	1	2	1	-	-	1	1	1	3
PK	-	3	2	1	-	1	-	1	6	1	-	1
Total	1	5	8	6	5	4	-	2	9	4	5	7

	BCDA rallentando 			**ADBC rallentando** 			**ABDC rallentando** 			**ACDB rallentando** 		
Speaker	R	P	S	R	P	S	R	P	S	R	P	S
AV	-	-	1	1	-	1	-	-	-	-	-	-
LG	-	-	-	-	-	1	-	-	-	1	1	1
OU	-	-	-	-	-	1	1	-	2	1	-	1
PK	-	-	-	-	-	-	-	-	1	1	-	-
Total	-	-	1	1	-	3	1	-	3	3	1	2

	CABD rallentando 			DABC rallentando 			CBDA accelerando 			CBAD accelerando 		
Speaker	R	P	S	R	P	S	R	P	S	R	P	S
AV	1	-	3	-	-	-	-	-	1	-	1	2
LG	1	1	1	-	-	-	-	-	1	2	-	-
OU	2	2	2	-	1	-	-	-	1	-	-	2
PK	1	-	-	-	-	-	-	-	-	1	1	1
Total	5	3	6	-	1	-	-	-	3	3	2	5

	CDBA accelerando 			ADCB accelerando 			BDCA accelerando 			CADB 		
Speaker	R	P	S	R	P	S	R	P	S	R	P	S
AV	-	-	-	-	-	-	-	-	-	-	-	-
LG	-	1	-	-	-	1	-	-	1	-	-	-
OU	-	-	-	-	1	-	-	-	-	1	-	-
PK	-	-	-	-	-	-	-	-	-	-	-	-
Total	-	1	-	-	1	1	-	-	1	1	-	-

	BADC 			BDAC 			CDAB 		
Speaker	R	P	S	R	P	S	R	P	S
AV	-	-	-	-	-	1	-	-	-
LG	1	-	-	-	1	-	-	-	-
OU	-	-	1	-	1	-	-	-	-
PK	1	-	-	-	-	1	-	-	1
Total	2	-	1	-	2	2	-	-	1

In grouping the occurring patterns into two more general patterns of rallentando and accelerando, the same criterion was used as in Chapter 2, i.e., at least three out of four component words within the phrase must follow the same direction of articulation rate change (down in the case of rallentando, and up in the case of accelerando). With this criterion in mind, the first ten patterns (ABCD, BACD, BCAD, ACBD, BCDA, ADBC, ABDC, ACDB, CABD and DABC) can be classed as rallentando, five patterns (CBDA, CBAD, CDBA, ADCB and BDCA)

as accelerando. The remaining patterns (CADB, BADC, BDAC and CDAB) do not seem to involve any unidirectional change of articulation rate.

In terms of the relative strength of these patterns, rallentando is clearly the dominant pattern. In reading, 78% of all 4-word phrases fall into this category, in paraphrasing 79%, and in spontaneous speech 75%. Accelerando is represented by 11% of cases in reading, 14% in paraphrasing, and by 17% in spontaneous speech. The very poor representation of the remaining patterns in all three tasks suggests that these patterns are exceptional.

Summary:

The results of the analysis of ordinal patterns are very much in line with the findings in Study 1, i.e., a progressive slowing down seems to be a typical pattern of articulation rate variation within the intonation phrase. It occurs consistently in all four speakers and in all three tasks. With regard to the actual distribution of individual patterns, it is worth noting how similar it is to the distribution found in Study 1. In particular, 3-word phrases exhibit again an overall balance between patterns ABC and BAC. In 4-word phrases, the more common patterns are those that can be seen as extensions of 3-word patterns, i.e., the same patterns as in 3-word phrases, but this time with the 4^{th} word (D) being the slowest.

CHAPTER 6

STUDY 2:

RESULTS OF THE STATISTICAL ANALYSIS

The analysis of ordinal patterns in the previous chapter seems to support Hypothesis 2.2, introduced in section 4.1; articulation rate manifests a tendency for a progressive slowing down throughout the intonation phrase and this tendency seems consistent both across speakers and across tasks. The present chapter tests this tendency statistically (Hypotheses 2.2 and 3.2, introduced in section 4.1), and it also deals with the issue of determinants of articulation rate variation (Hypotheses 4.2-7.2 from section 4.1). All the hypotheses were examined using a GLM analysis.

6.1 ASSUMPTIONS FOR GLM ANALYSIS

As described in section 3.4.1, there are three basic assumptions underlying the use of the analysis of variance: normal distribution of observations on the dependent variable within each group, homogeneity of variance, and independence of observations one from another.

The normality of the distribution was examined by using the Kolmogorov-Smirnov test and Normal Q-Q plots. The Kolmogorov-Smirnov test suggested that in some groups the assumption of normality of distribution was not satisfied. However, as noted in section 3.4.1, the Kolmogorov-Smirnov test is a sensitive test and it can be affected even by a small number of outliers while the distribution is, in fact, normal. To make sure that this is the case for my material, I examined also Normal Q-Q plots. They did not show any considerable departures

from normality. The assumption of normal distribution was therefore considered satisfied.

Homogeneity of variance was tested by the Levene test, which indicated some differences in variance across groups, but examination of boxplots suggested that the differences are due to few outliers and are, in fact, rather small. Since the analysis of variance is known to be fairly robust against a moderate degree of heterogeneity in variances, the assumption about the homogeneity of variance can also be considered satisfied in the material.

The third assumption, the independence of observations one from another, was satisfied as well. In all three tasks the subjects produced a continuous speech (i.e., no repetitions) under the same conditions. As in Study 1, there was nothing in the experimental setup which would induce subjects to make the articulation rate of one word influence the articulation rate of another.

6.2 OVERALL MODEL

Testing Hypotheses 2.2-7.2 involved two stages of analysis. In the first stage, a model with the main factors and 2-way and 3-way interactions was used.

Examining interactions generally shows whether the factors are independent one from another. In the present study they were examined mainly to test those parts of Hypotheses 3.2, 4.2 and 5.2 which refer to the comparison of speakers and tasks. Thus, for instance, a 2-way interaction position*speaker reveals whether there are any systematic differences in the patterns of articulation rate variation across speakers; a 2-way interaction position*task indicates possible different patterns across tasks. But the other interactions are also of interest. For instance, a 2-way interaction position*word class tests whether the mean articulation rate of words in different positions within the intonation phrase is significantly affected

by the word class. Interactions of a higher order than the second (i.e., 3-way interactions) are difficult to interpret since there is no obvious linguistic mechanism which would account for their plausibility. They are expected to be non-significant. Thus a saturated model (i.e., a model incorporating all main factors and all possible interactions) was not performed; the second order (i.e., 3-way) interactions were the highest order considered. The second stage, the core analysis, involved modelling articulation rate only with the significant main factors and those interactions which were significant at the first stage.

For both stages of the analysis, only the most relevant data are presented in the main body of the text; full data can be found in Appendices. For the first-stage models only the information about which factors and interactions were significant and at what level will be mentioned in the text; further information, namely, sums of squares, degrees of freedom and F ratios are included in Appendix 6. Regarding the main analysis, more extensive results will be reported in the text, particularly, F ratios, degrees of freedom and levels of significance. The remaining details of F statistics can be found in Appendix 7. Appendix 7 also provides full details on Tukey tests (mean differences, standard errors and significance levels); the text includes only the most important information, i.e., which levels within the factors were significantly different from each other. As mentioned in section 4.8, 2-word, 3-word and 4-word intonation phrases were analyzed separately.

6.2.1 Model with the main factors, 2-way and 3-way interactions

The GLM analysis on the main factors (position, word size, word class, speaker and task), and all 2-way and 3-way interactions yielded the following results. Position and word size were highly significant in all three groups of phrases (both factors with $p < 0.001$ in 2-word and 3-word phrases; with $p = 0.004$ and $p = 0.006$ respectively in 4-word phrases). Word class and speaker were significant

only in 3-word phrases ($p < 0.001$ and $p = 0.012$ respectively). Task was non-significant in all three groups of phrases ($p > 0.1$).

No 2-way interactions were significant in 2-word and 4-word intonation phrases ($p > 0.05$); in 3-word phrases only the 2-way interaction position*speaker proved significant ($p = 0.021$). Of the 3-way interactions only position*word size*word class was found significant in 2-word phrases (rather weakly, $p = 0.036$) and 4-word phrases ($p = 0.01$), and word size*word class*task in 3-word phrases ($p = 0.045$).

6.2.2 Results of the main analysis

In the core GLM analysis, all the main factors were included in the model although some of them were non-significant when combined with 2-way and 3-way interactions. The non-significance may be a side-effect caused by the nature of the GLM procedure; it is possible that the result is spurious due to the large number of interactions in the model, and after removing the non-significant interactions from the model the main factors may well become significant. (This proved to be the case in my study, as will be shown below.) Moreover, the picture was not consistent across the three groups of phrases and thus including all main factors in the core analysis helps to clarify the situation.

In the GLM for 3-word phrases, a 2-way interaction position*speaker was also incorporated in the model, since it proved significant at the first stage of the analysis. This indicated that differences among the speakers as to the patterns of articulation rate variation might occur. However, it is also possible that, as for some of the main factors, this result is spurious due to the large number of interactions in the model. In order to obtain a clearer picture, mean articulation rates in different word positions within the individual speakers will be discussed at a later stage of the analysis.

Although at the first stage of the analysis a 3-way interaction position*word size*word class turned out to be significant in some cases, it was not included in the main analysis. The reasons are, firstly, the possibility of a spurious result as a consequence of a large number of interactions, as described above, and, secondly, the difficulty interpreting multi-way interactions.

Table 6.1 (with full results reported in Appendix 7) shows that all main factors were highly significant, apart from task in 2-word phrases ($p > 0.1$), and in 3-word phrases, in which its significance level was weaker ($p = 0.017$). Since task was non-significant in 2-word phrases, the GLM was re-run for this group of phrases, this time without task. In this analysis, all the remaining main factors were highly significant. The 2-way interaction position*speaker in 3-word phrases proved highly significant.

Table 6.1 also includes information about the overall success of the models. Both the coefficient of determination, R^2, and the adjusted R^2 are presented for this purpose. The adjusted R^2 allows comparison between 2-word, 3-word and 4-word phrases, since the three models incorporate a different number of factors (see the discussion in section 4.8). When the models are considered separately, using R^2 is sufficient. This measure suggests that modelling articulation rate was reasonably successful. In 2-word phrases, 53% of variance in articulation rate was accounted for; in 3-word phrases, 52%, and in 4-word phrases, 55%.

To provide a rough estimate of the strength of the individual factors within the sample, a measure of the strength of association, η^2, was calculated. Recall from section 4.8 that the formula used for the calculation is as follows:

$$\eta^2 = SS_x / SS_{Total}$$

The results are summarized also in Table 6.1.

Table 6.1 **Results of the GLM analysis for 2-word, 3-word and 4-word intonation phrases.**

2-word IP's	*df*	*F*	*p*	R^2	*Adj.* R^2	η^2
Model	**11,718**	**74.7**	**< 0.001**	**53.4%**	**52.7%**	
Position	1,718	192.3	< 0.001			0.125
Word size	5,718	52.2	< 0.001			0.169
Word class	2,718	35.9	< 0.001			0.047
Speaker	3,718	6.5	< 0.001			0.017
3-word IP's						
Model	**20,903**	**48.6**	**< 0.001**	**51.8%**	**50.8%**	
Position	2,903	141.7	< 0.001			0.151
Word size	5,903	42.8	< 0.001			0.114
Word class	2,903	67.7	< 0.001			0.072
Speaker	3,903	13.7	< 0.001			0.022
Task	2,903	4.1	= 0.017			0.004
Position*Speaker	6,903	4.9	< 0.001			0.016
4-word IP's						
Model	**15,432**	**34.6**	**< 0.001**	**54.5%**	**53%**	
Position	3,432	45.0	< 0.001			0.142
Word size	5,432	33.9	< 0.001			0.178
Word class	2,432	29.9	< 0.001			0.063
Speaker	3,432	10.3	< 0.001			0.033
Task	2,432	7.0	= 0.001			0.015

The measure of the strength of association indicates that the factors affecting the word's articulation rate to the greatest extent were the word's position within the phrase and the number of syllables it contains. In comparison with these, the factors word class, speaker and task were weaker, with task being at the weakest point of the scale. In 2-word and 4-word phrases, word size came out as stronger than position. However, the values for η^2 are relatively close and no definite conclusion about the respective strengths of these two factors can be drawn.

Although the GLM analysis established that position, size, word class, speaker (and possibly) task are important factors affecting the words' articulation rate, it did not reveal the nature of the relationship between the word's articulation rate and these factors. Therefore further analysis involved examination of the means in some cases, boxplots and conducting the Tukey test. I shall discuss the findings for all the factors in turn.

6.2.2.1 Position

The nature of the relationship between position of the word and its articulation rate, i.e., a progressive slowing down throughout the intonation phrase, is the subject of Hypothesis 2.2. Table 6.2 below presents the mean articulation rates and standard deviations for words in different positions within the phrase. The table also provides the information, obtained from the Tukey test, about which positions were significantly different from each other. The combined examination of means and of the post-hoc tests addresses the question of the possible linearity of the slowing down (Hypothesis 3.2).

The findings generally confirmed the observation made on the basis of ordinal patterns (section 5.3) that articulation rate within the intonation phrase tends to reflect a pattern of slowing down. In 2-word phrases, the slowing down was necessarily linear (the overall pattern AB) and the mean articulation rates in the two positions were significantly different, as indicated by the significance of the factor position in the GLM analysis (Table 6.1).

Table 6.2 **Mean articulation rates [sylls/s] and standard deviations (SD) summed across speakers and tasks, for words in different positions within the intonation phrase, and the results of the Tukey test ($p < 0.05$) on which positions involved significantly different means. (Positions are represented by numbers, e.g., 1-2 indicates that the mean articulation rates in the positions 1 and 2 within the phrase were significantly different.)**

	2-word IP's		3-word IP's		4-word IP's	
	Mean	SD	Mean	SD	Mean	SD
Position 1	6.97	1.74	7.45	1.84	7.34	1.90
Position 2	4.81	1.40	6.95	1.80	6.93	1.85
Position 3			4.79	1.35	6.56	1.59
Position 4					4.73	1.39
Significantly different means (positions)	1-2		1-2 1-3 2-3		1-3 1-4 2-4 3-4	

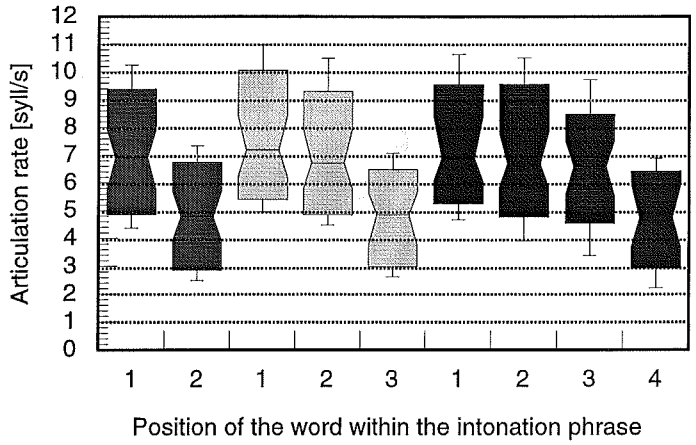
In 3-word phrases, the slowing down (the overall pattern ABC) was non-linear, with the first two words having their mean articulation rates closer together than they have in relation to the last word. (Recall from section 3.4 that this is scenario 2a.) The Tukey test showed nonetheless that the first two words were significantly different. With regard to the third, final word, it is hardly surprising that it is significantly different from the first two words, given that its mean rate is clearly much lower than their mean rates.

In 4-word phrases (the overall pattern ABCD) the slowing down was also non-linear, with the mean articulation rate of the last word within the phrase being significantly lower than the rate of all the other words within the phrase. This is in line with the expectation that phrase-final lengthening applies to the intonation phrase. The mean articulation rates of the first two words were not statistically different from each other and neither were the rates of the second and the third words. However, the fact that the first and the third words were also statistically different suggests that final lengthening is not the only factor in play and that

some amount of a progressive slowing down does occur in 4-word phrases as well.

The non-linearity of the slowing down in 3-word and 4-word phrase is illustrated in Figure 6.1, which contains boxplots for factor position. Boxplots used in the present study contain the following information: the median, the 25th and 75th percentiles (notches), the 10th and 90th percentiles (boxes), and the 5th and 95th percentiles (whiskers). Extreme cases were not included in the display. The boxplots are always followed by the information about the number of observations within categories.

Figure 6.1 **Boxplot for articulation rate in different positions within the intonation phrase.**
Dark grey: 2-word phrases; light grey: 3-word phrases; black: 4-word phrases.



Number of words in different positions within the intonation phrase
2-word IP's - 365 in each position 3-word IP's - 308 in each position 4-word IP's - 112 in each position

A regular pattern of slowing down within the intonation phrase was confirmed as consistent across tasks in all three groups of intonation phrases (the 2-way interaction position*task was non-significant in the first-stage model). It was also confirmed as consistent across speakers for 2-word and 4-word phrases. However, in 3-word phrases possible differences were indicated as to what pattern of rate variation was typical for different speakers (recall that the 2-way interaction position*speaker was highly significant in this group of phrases). Further analysis was thus conducted to find out whether such differences are genuine, and if so, what their nature is.

In this analysis, the overall data set for 3-word phrases was split into 4 subsets, one for each speaker, with task remaining combined within each set. The means, standard deviations and the Tukey test on the significance of differences between means were examined. The results are displayed in Table 6.3 and in Figure 6.2 (full details can be found in Appendix 8).

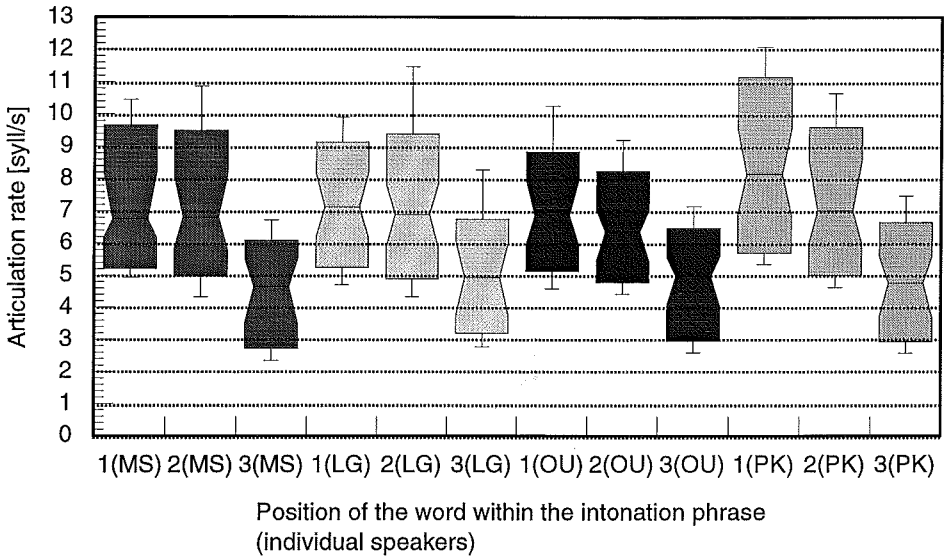
Table 6.3 **Mean articulation rates [sylls/s] and standard deviations (SD) for words in different positions within the intonation phrase: the results for individual speakers in 3-word phrases (tasks combined). The results of the Tukey tests ($p < 0.05$) on which positions involved significantly different means.**

	AV		LG		OU		PK	
	mean	SD	mean	SD	mean	SD	mean	SD
Position 1	7.19	1.80	7.12	1.54	6.99	1.54	8.34	2.09
Position 2	7.06	1.78	7.08	2.06	6.41	1.45	7.26	1.76
Position 3	4.54	1.29	5.00	1.47	4.83	1.25	4.73	1.37
Significantly different means (positions)	1-3 2-3		1-3 2-3		1-2 1-3 2-3		1-2 1-3 2-3	

The analysis revealed that the patterns of rate variation in individual speakers are not, in fact, conflicting in any strong sense. The overall pattern is ABC for all four

speakers, that is, the same as for the set of data with all speakers combined. The difference seems to be in the degree of linearity of the slowing down. For two speakers, OU and, particularly, for PK, it is noticeably more linear than for speakers AV and LG. This is reflected in the results of the Tukey test. It showed that in samples from OU and PK, all words were significantly different from each other in their mean rate, while in samples AV and LG the first two words were not (in fact, their means were nearly identical).

Figure 6.2 **Boxplot for articulation rate in different positions within the intonation phrase: individual speakers (in brackets).**



A comparison of the frequencies of the ordinal patterns ABC and BAC for the individual speakers (section 5.3.2) suggests that the reason might lie partly in a stronger prevalence of the pattern ABC for speakers OU and PK in nearly all tasks, as opposed to the more balanced frequencies of the two patterns for speakers AV and LG. (Summed across all the three tasks, the frequency distribution of ABC and BAC patterns, i.e., the number of phrases with these

patterns, was as follows: subject AV: 29 vs. 24; subject LG: 25 vs. 28; subject OU: 37 vs. 26; subject PK: 38 vs. 31.)

In the analysis of 3-word phrases in Study 1, the nearly identical mean articulation rates for words in positions 1 and 2 also reflected an overall balanced number of occurrences of ABC and BAC patterns. However, the separate statistical analyses of these two patterns in section 3.4.3 provided further support for the claim that ABC pattern is a default pattern of articulation rate variation within the intonation phrase. In the present study, typicality of pattern ABC has already been demonstrated in the above statistical analysis. However, a similar analysis to that in Study 1 was also carried out for the present material to find out whether similar results emerge. I shall report on this analysis after discussing the remaining factors in the overall GLM analysis.

6.2.2.2 Word size

The factor of word size proved to be the strongest influence on the word's articulation rate, alongside position. The investigation of the nature of the relationship between articulation rate and word size again involved conducting the Tukey test and examining the boxplots (Figure 6.3) for groups of words containing a particular number of syllables.

The tendency is clear and, as shown in Table 6.1, highly significant; in all three groups of phrases (2-word, 3-word and 4-word) the articulation rate increases with the number of syllables in a word. However, Figure 6.3 illustrates that the trend is not linear. This is reflected in the results of the Tukey test, which revealed that not all the categories of syllable size are equally responsible for the overall significant result. The details about which categories were significantly different from each other in their mean articulation rates are shown in Table 6.4. In order to be able to perform the Tukey test on 4-word phrases, the category of 6 syllable words had to

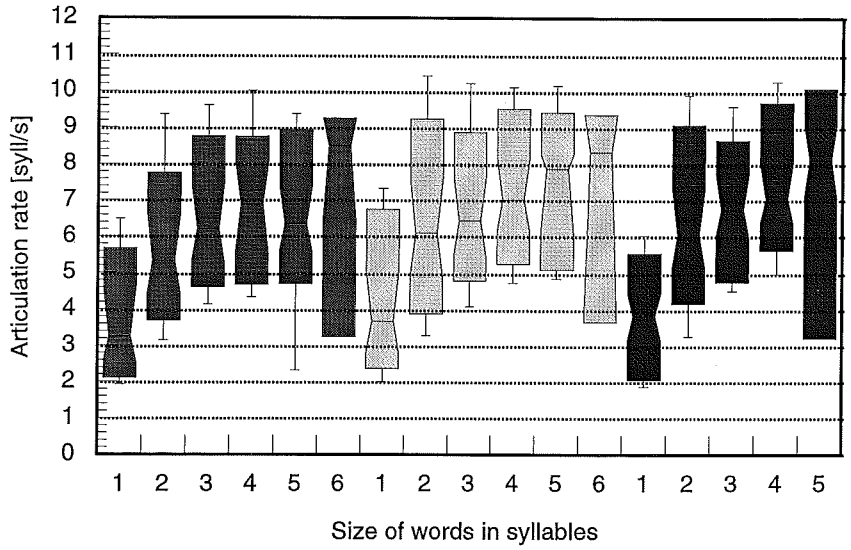
be excluded, since the calculation of the Tukey test requires that each category involves more than one observation, which was not the case for this category.

The most distinct word size category in all three groups of phrases were monosyllables (the mean articulation rate 3.83 syll/s when the three groups of phrases are combined).[1] The mean articulation rate of monosyllabic words was significantly lower than the rate of all the other word size categories, including disyllables (6.14 syll/s). The difference between disyllabic and trisyllabic words (whose overall mean was 6.65 syll/s) was smaller than between monosyllables and disyllables and yet it was significant. Disyllables were also significantly slower than four-syllable words in all three groups of phrases (after combining them, 7.07 syll/s), and in case of 3-word and 4-word phrases, also slower than five-syllable words (7.51 and 8.13 syll/s respectively). The differences among the words containing three or more syllables were all non-significant. However, in order to be able to draw stronger conclusions about the variation in articulation rate across different word sizes, a more balanced set of data across the categories would be needed, in particular, more 5- and 6-syllable words.

The results apply to all positions within the phrase (otherwise, the 2-way interaction position*word size would have been significant), whether the word is a content word or function word (the 2-way interaction word category*word size was non-significant). The results were also consistent across tasks and speakers.

[1] Combining the three groups together to obtain overall means for individual categories of word size was possible since the categories are of the same kind across 2-, 3- and 4-word phrases. This was not the case for factor of position.

Figure 6.3 **Boxplot for articulation rate in words with different number of syllables.**
Dark grey: 2-word phrases; light grey: 3-word phrases; black: 4-word phrases.



Number of words of different size			
Word size [sylls]	2-word IP's	3-word IP's	4-word IP's
1	86	72	43
2	288	433	202
3	234	292	153
4	103	106	41
5	16	19	8
6	3	2	(1)

Table 6.4 **Categories of words with different size in syllables which were significantly different from each other.**
(The results of the Tukey test.)

2-word IP's	3-word IP's	4-word IP's
1-2 2-3	1-2	1-2 2-3
1-3 2-4	1-3 2-4	1-3 2-4
1-4	1-4 2-5	1-4 2-5
1-5	1-5	1-5
1-6	1-6	

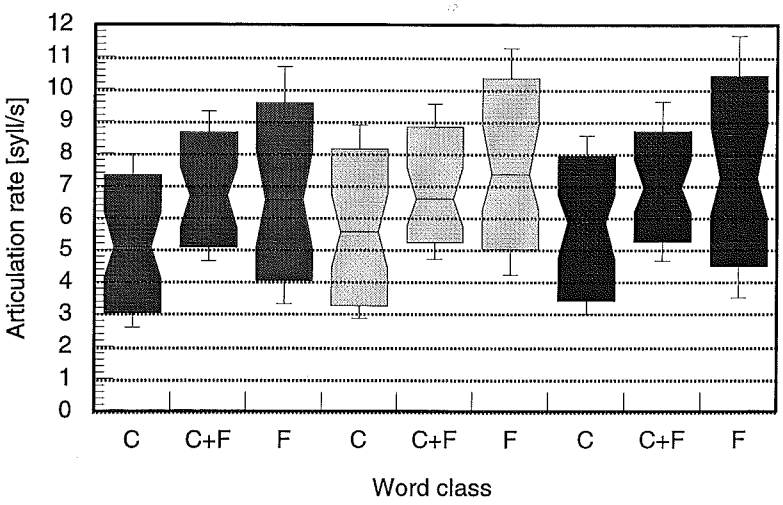
6.2.2.3 Word class

It was shown in Table 6.1 that the factor of word class significantly affected the articulation rate of words, although to a noticeably smaller extent than position and word size. The boxplots in Figure 6.4 and the Tukey test (Table 6.5) show that phonological words containing only content word(s) tended to be significantly slower than those containing content word plus one or more function words, or function word(s) only. This applies to all three groups of phrases. The phonological words incorporating only function word(s) tended to be the fastest of the three categories in 3-word and 4-word phrases. However, they were significantly faster than content+function word category only in 3-word phrases.

The tendency of the function word category to be the fastest occurred despite the fact that over 50% of the phonological words concerned consisted of two or three, rarely four, lexemes (the remaining part being single lexemes).[2] On the other hand, the absolute majority of phonological words classified as being in the content word category were single lexemes. The present findings suggest that the function vs. content word opposition is indeed an important factor affecting the word's articulation rate; if the unit in the study were lexemes (as often they have been, e.g., O'Shaughnessy, 1995) the difference between content and function words would probably be even more pronounced.

[2] This distribution is a consequence of the tendency of function words to loose their stress in connected speech and cliticize.

Figure 6.4 **Boxplot for articulation rate in words of different word class. Dark grey: 2-word phrases; light grey: 3-word phrases; black: 4-word phrases.**



Legend: C = phonological word consisting of one or more content words
C+F = phonological word containing a content word and one or more function words
F = phonological word consisting of one or more function words

Number of words of different word class			
Word class	2-word IP's	3-word IP's	4-word IP's
C	400	494	262
C+F	148	177	73
F	182	253	113

Table 6.5 **Word class categories which were significantly different from each other. (The results of the Tukey test.)**

2-word IP's	3-word IP's	4-word IP's
C - F C - C+F	C - F C - C+F F - C+F	C - F C - C+F

In 2-word phrases the rate of the content+function word category was faster than that of the function word category, i.e., the reverse of the trend in 3-word and 4-word phrases. As shown, the difference between the two categories was not significant. It is an open question whether the trend would become more similar to 3- and 4-word phrases, if even a larger sample size was available.

One could argue that the word class classification used in this study, especially separating out the content+function word category, is rather arbitrary. Yet the finding that in 3-word phrases (incorporating nearly half of the total number of words) the categories were significantly different from each other indicates that they all do play a role in determining the articulation rate of phonological words. As claimed above, the most important distinction is between content and function words. However, no generalizations about the relationship between the function and the content+function word categories can be made on the basis of the present material.

The tendency for the function word category to have the fastest articulation rate, and the content word category to be the slowest, was consistent across speakers and tasks. The word class categories behaved in the same way also in different positions within the phrase and irrespective of the number of syllables they contain. These conclusions follow from the fact that in the model with 2-way and 3-way interactions (section 6.2.1 and Appendix 6) the 2-way interactions between word class and the above factors respectively were not significant.

6.2.2.4 Relationships between position, word size and word class

The relationship between the word class categories and the number of lexemes in a phonological word, which I mentioned at the beginning of section 6.2.2.3, leads to the question whether there might also be some relationship between word class and word size. The two factors might have to some extent a confounding effect on

articulation rate. In fact, the factor of position in combination with word size and word class could also have a similar effect. For instance, one might expect a certain word class to occur usually in a certain position within the phrase. The study of articulation rate involving the GLM analysis does not become seriously affected by confounding effects, since the model would reveal completely confounding effects by manifesting a non-significant effect for one of the factors involved. Yet it is of interest to find out more about the relationship between position, word size and word class.

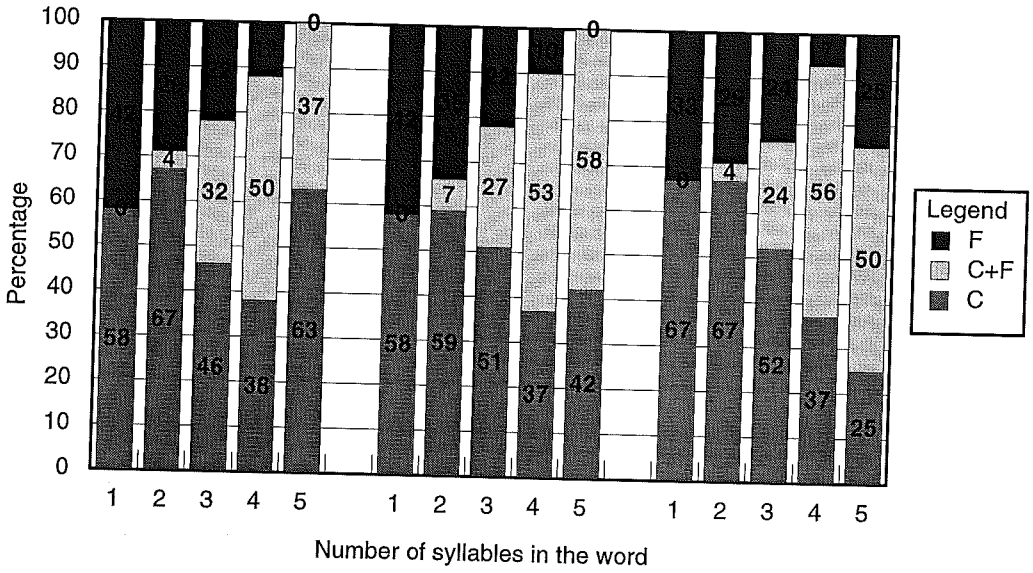
First of all, I shall discuss the relationship between word class and word size. In general, function words tend to be monosyllabic. The present study does not work with lexemes but phonological words, and thus the word class categories include also phonological words consisting of several lexemes. Those words which were classified as belonging to the function word category contained a single lexeme - function word - in only about 50% of cases. The other 50% contained more than one function word and therefore must have contained more than one syllable. A general trend was that the function word category had a significantly faster articulation rate than the content word category. The trend for word size was one of increasing rate with increasing number of syllables in the word. Thus if there is an interaction between the number of syllables in the word and the word class with respect to articulation rate, the function word category should incorporate mostly polysyllabic words. In such a case, the association of the fastest articulation rate with the function word category could be partly due to the larger number of syllables in this category. This, however, does not seem to be the case. A representation of the relationship between word size and word class is shown in Figure 6.5.

Around 60% of monosyllabic and disyllabic words belong to the content word category, the rest to the function word category. In 2-word and 3-word phrases there was only a small percentage of four-syllable representatives of the function word category and none of five-syllable ones. However, rather few of these polysyllabic words belong to the content word category; mostly they are in the

content+function word category. Altogether, these observations suggest that the tendency for the content word category to be the slowest one and the function word category to be the fastest is unlikely to be due to any interaction with word size.[3]

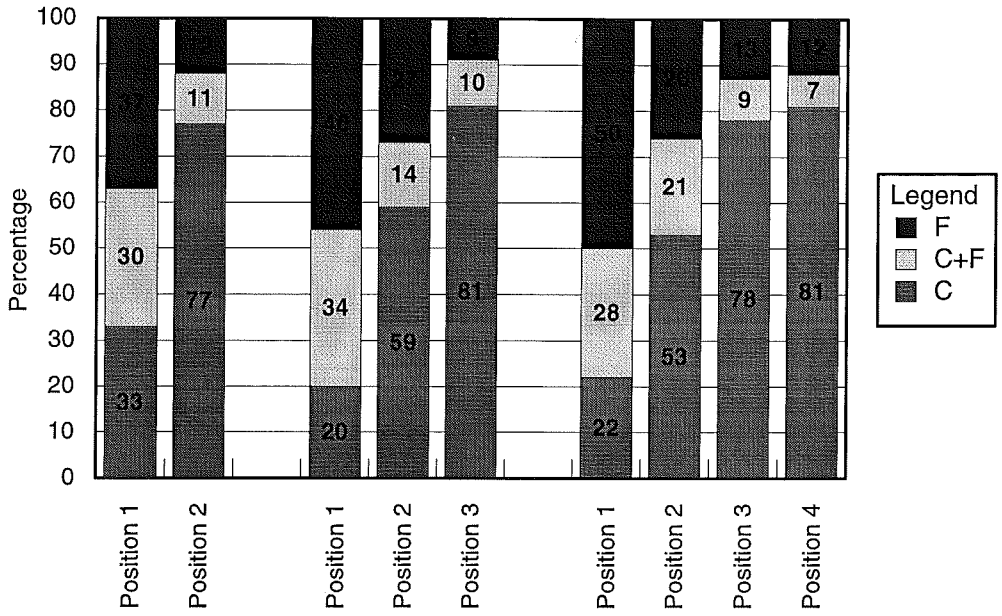
Figure 6.5 **Distribution of categories of word class for words consisting of different number of syllables (groups of five columns in the order of 2-word, 3-word and 4-word phrases).**

C: content only; C+F: content + function; F: function only



[3] It should be noted that the number of observations within the categories is unbalanced (see the data within Figure 6.4) and thus the tendencies described here should be taken as tentative.

Figure 6.6 **Distribution of word class categories in different positions within the intonation phrase.**



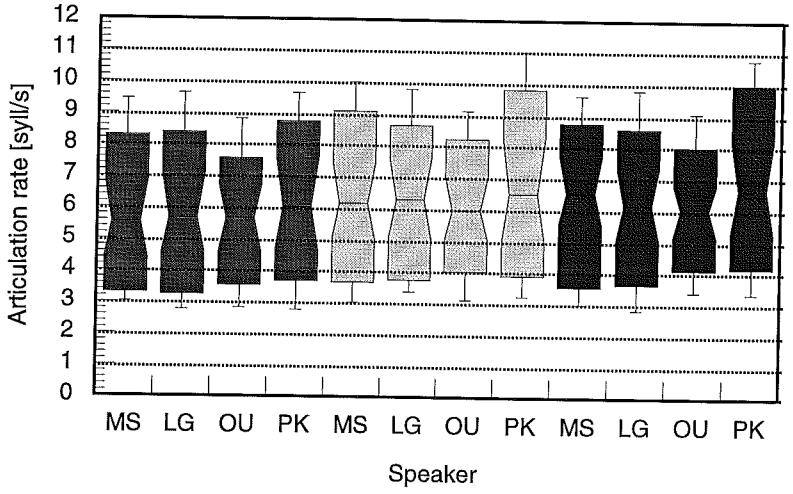
The relationship between the position of the word within the phrase and the word class is more straightforward. As seen from Figure 6.6, the function word category tends to be in the first position while the content word category prevails in the later positions within the phrase, with the absolute majority occupying the last position. This suggests a possible interaction between position and word class with respect to articulation rate. This raises the question whether position and word class are correlated to such an extent that position could be replaced by word class in the GLM analysis, rather than the two being treated as separate factors. The GLM procedure demonstrated a highly significant effect on articulation rate for both factors; unfortunately, this does not guarantee an absence of a confounding effect. Looking at Figure 6.6, we can see that, although content words clearly prevail in the final position within the phrase, function words do not prevail in the initial position to such a large extent. They take up no more than 50% of the total number of words in this position, while content words take up about 80% of words in the final position in all three groups of phrases. Since the interaction between position and word class is an important issue, it deserves

further investigation. Such an investigation would need an advanced statistical processing of the kind which would be able to resolve the issue of confounding effects. However, such processing was not available to me.

6.2.2.5 Speaker

In Hypothesis 6.2 the factor of speaker was predicted to significantly affect a word's articulation rate. Table 6.1 in section 6.2.2 confirmed this hypothesis. The measure of the strength of association indicated that although this factor was highly statistically significant, its power is rather weak in comparison with word class and, in particular, with position and word size. Boxplots in Figure 6.7 show the differences between the speakers, and Table 6.6 presents the results of the Tukey test on which of these differences were statistically significant.

Figure 6.7 **Boxplot for articulation rate in different speakers. Dark grey: 2-word phrases; light grey: 3-word phrases; black: 4-word phrases.**



Number of words for different speakers			
Speaker	2-word IP's	3-word IP's	4-word IP's
MS	170	189	104
LG	172	237	108
OU	178	237	136
PK	210	261	100

Table 6.6 **Speakers who were significantly different from each other. (The results of the Tukey test.)**

2-word IP's	3-word IP's	4-word IP's
MS - PK OU - PK	MS - PK OU - PK	LG - PK OU - PK

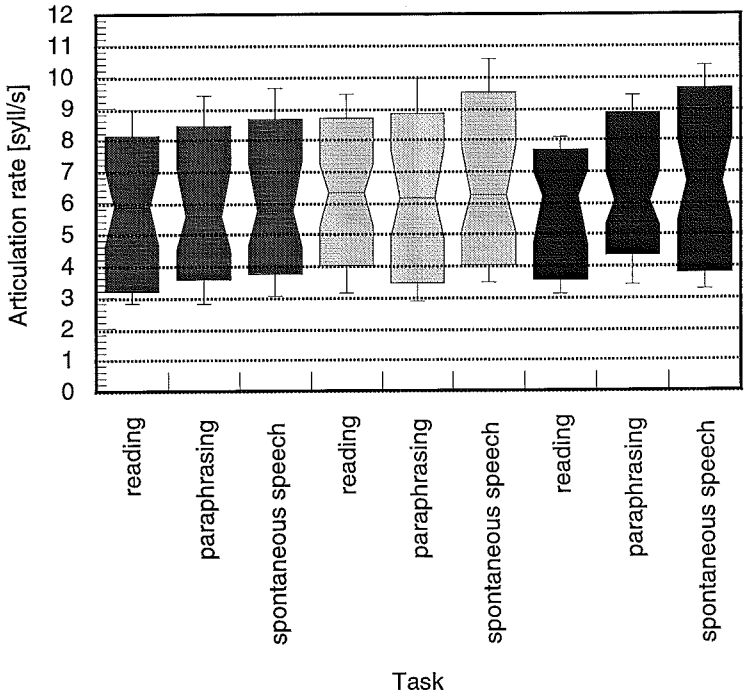
The results of the analysis revealed that not all speakers were significantly different from each other in their mean articulation rate. The one who seems to be responsible for the overall high significance of this factor is speaker PK, who has the fastest articulation rate of all speakers in all three groups of phrases (when all three groups of phrases combined, his mean articulation rate was 6.6 syll/s). However, even this speaker was not *always* significantly different from the other three speakers. He was consistently faster than speaker OU only (5.94 syll/s). Speakers AV (6.12 syll/s), LG (6.19 syll/s) and OU were not significantly different from each other.

6.2.2.6 Task

The last main factor investigated, task, proved in the GLM analysis to be the weakest factor affecting the articulation rate of words; it was highly significant only in 4-word phrases, somewhat less significant in 3-word phrases and not at all in 2-word phrases (Table 6.1). The prediction formulated in Hypothesis 7.2 was

that reading in the present material would be significantly faster than spontaneous speech, and this would be, in turn, significantly faster than paraphrasing. However, the examination of the mean articulation rates in the three tasks did not confirm this hypothesis. In all three groups of phrases, the fastest articulation rate was found in spontaneous speech (2-word phrases: 6.01 syll/s; 3-word phrases: 6.56 syll/s; 4-word phrases: 6.69 syll/s). Reading and paraphrasing were nearly identical in their mean articulation rates in 2-word and 3-word phrases (2-word phrases: 5.83 syll/s and 5.81 syll/s respectively; 3-word phrases: 6.32 syll/s and 6.26 syll/s respectively). In 4-word phrases, reading manifested a slower articulation rate than paraphrasing (5.89 syll/s, as opposed to 6.27 syll/s). However, the Tukey test showed that the three tasks were not significantly different from each other in 2-word phrases; in 3-word phrases, only spontaneous speech was significantly faster than paraphrasing, and in 4-word phrases, a significant difference occurred only between spontaneous speech and reading.

Figure 6.8 **Boxplot for articulation rate in different tasks. Dark grey: 2-word phrases; light grey: 3-word phrases; black: 4-word phrases.**



Number of words in different tasks			
Task	2-word IP's	3-word IP's	4-word IP's
Reading	182	273	108
Paraphrasing	266	282	112
Spontaneous speech	282	369	228

It is also worth noting that examination of the means indicated a somewhat different order of the tasks with respect to articulation rate than examination of the medians in the boxplots of Figure 6.8. The differences between means and medians are likely to be due to some outliers.[4] The boxplots in Figure 6.8 suggest that reading was marginally faster than both spontaneous speech and paraphrasing in 2-word and 3-word phrases; in 4-word phrases, the fastest task was clearly spontaneous speech.

Non-significant results for two-way interactions speaker*task indicate that the four speakers did not differ with respect to which tasks they performed at a faster and which at a slower rate.

6.2.2.7 Patterns ABC and BAC

I shall come back now to patterns ABC and BAC. Separate GLM analyses were run on each pattern for individual speakers (but not individual tasks, since the interactions task*position and task*position*speaker were non-significant in the first stage of the main analysis). The factors were position of the word, its size, word class, and task; the dependent variable was, of course, articulation rate.

The factor which was most important in this analysis was position. It was highly significant in all four speakers in both patterns ABC and BAC (ABC: for AV, [*F*

[4] The issue of a different order of levels within a category did not arise for any of the other main factors.

(2,74) = 29.82, $p < 0.001$], for LG [F (2,65) = 47.89, $p < 0.001$], for OU, [F (2,100) = 45.80, $p < 0.001$], and for PK, [F (2,106) = 59.62, $p < 0.001$]; BAC: for AV, [F (2,60) = 45.88, $p < 0.001$], for LG [F (2,74) = 29.04, $p < 0.001$], for OU, [F (2,67) = 18.49, $p < 0.001$], and for PK, [F (2,82) = 39.15, $p < 0.001$]).

Word size also affected the articulation rate of word highly significantly in both patterns ABC and BAC, although in the case of LG's BAC pattern the relationship between word size and its articulation rate was much weaker.[5]

Word class was generally a much weaker factor. In pattern ABC, it was highly significant only in two speakers; in pattern BAC, the significance was lower overall or absent.[6] Finally, task was very weakly significant for two speakers in pattern ABC and for no speakers in pattern BAC.[7]

Table 6.7 shows the mean values for articulation rate in different positions within the phrase, separately for pattern ABC and BAC.

Unlike in Study 1, where the first two positions in pattern BAC were not significantly different in their mean articulation rates, in this study, both patterns ABC and BAC had all the positions significantly different from each other. This suggests that when analysed on its own, even pattern BAC has all the positions distinct in terms of articulation rate, if there is a sufficiently large number of samples.

---

[5] Pattern ABC: for speaker AV, [F (3,74) = 5.34, p = 0.002], for LG [F (3,65) = 4.77, p = 0.005], for OU, [F (4,100) = 7.42, p = 0.001], and for PK, [F (4,106) = 4.19, p = 0.003]. Pattern BAC: for AV, [F (5,60) = 3.70, p = 0.005], for LG [F (3,74) = 3.12, p = 0.031], for OU, [F (4,67) = 3.67, p = 0.009], and for PK, [F (4,82) = 4.44, p = 0.003].

[6] Pattern ABC: for speaker AV, [F (2,74) = 4.75, p = 0.048], for LG [F (2,65) = 6.75, p = 0.002], for OU, [F (2,100) = 4.62, p = 0.012], and for PK, [F (2,106) = 5.23, p = 0.007]. Pattern BAC: for AV, [F (2,60) = 4.78, p = 0.012], for LG [F (2,74) = 3.91, p = 0.024], for OU, [ns, $p >$ 0.1], and for PK, [F (2,82) = 4.42, p = 0.015].

[7] Pattern ABC: for speaker LG [F (2,65) = 3.68, p = 0.031], and for PK, [F (2,106) = 3.31, p = 0.04]. For speakers AV and OU in ABC pattern, and for all speakers in BAC pattern, task was non-significant ($p > 0.1$).

Table 6.7 **Patterns ABC and BAC: Mean articulation rates [sylls/s] and standard deviations for words in different positions within the intonation phrase, and the results of the Tukey test on which positions involved significantly different means ($p < 0.05$).**

ABC 	Speaker AV		Speaker LG		Speaker OU		Speaker PK	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Position 1	7.90	1.73	7.84	1.12	7.79	1.32	9.47	1.87
Position 2	6.52	1.16	6.49	0.74	6.06	0.84	6.89	1.51
Position 3	4.38	0.99	4.43	1.02	4.67	1.15	4.36	1.13
Significantly different means (positions)	1-2 1-3 2-3		1-2 1-3 2-3		1-2 1-3 2-3		1-2 1-3 2-3	

BAC 	Speaker AV		Speaker LG		Speaker OU		Speaker PK	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Position 1	6.47	1.17	6.89	1.15	6.24	0.93	7.13	1.21
Position 2	8.34	1.67	8.40	2.02	7.34	1.60	8.25	1.42
Position 3	4.27	1.51	4.44	1.28	4.35	1.07	4.46	1.34
Significantly different means (positions)	1-2 1-3 2-3		1-2 1-3 2-3		1-2 1-3 2-3		1-2 1-3 2-3	

In the discussion of patterns ABC and BAC in Study 1 (section 3.4.3), I brought up the question of how these two patterns relate to the linguistic structure of relevant phrases and whether this relationship might explain the particular distribution of these two patterns. One aspect of linguistic structure to look at again is the size of words in syllables. A highly significant tendency for the word's articulation rate to increase with the number of syllables in the word was found in the present study, even when patterns ABC and BAC were analyzed separately (recall that in Study 1, this factor was not significant at this stage of analysis). Therefore, I examined whether in pattern ABC mostly polysyllabic words occurred in the first position (position A), while in pattern BAC, in the second position (position B). Figure 6.9 presents the distribution of words

according to their size in syllables for different positions within the intonation phrase.

Figure 6.9 **Distribution of words according to their size in syllables for different positions within the intonation phrase.**

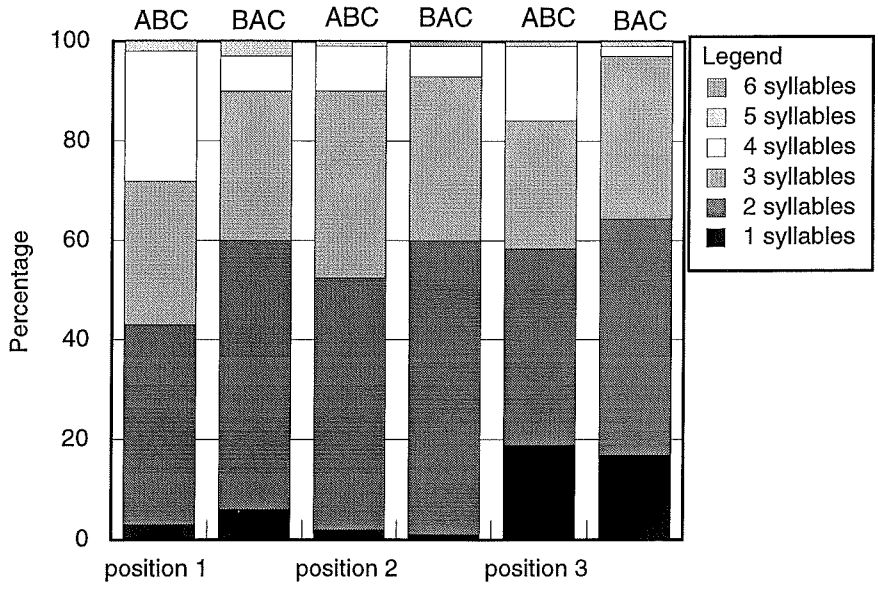
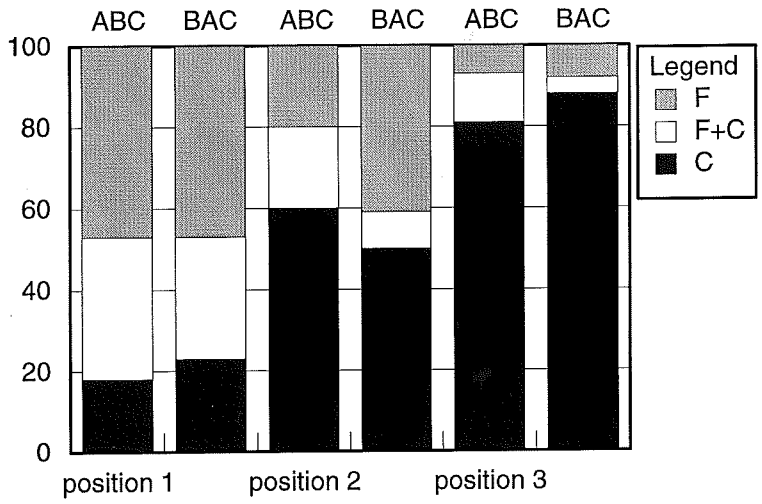


Figure 6.9 indicates that the conjectured distribution did not fully apply. In the first position of pattern ABC, there are comparatively more words with three and four syllables than in pattern BAC (55%, as opposed to 37%). Pattern BAC involves in the first position mostly monosyllabic and disyllabic words, unlike pattern ABC (60%, as opposed to 43% in the ABC pattern). This might contribute to a slower articulation rate in the first position in the BAC pattern than in the ABC pattern. However, when the second position was compared in patterns ABC and BAC, no tendency was found for polysyllabic words to prevail in the BAC pattern. Neither were shorter words found to prevail in the second position of the ABC pattern. The distribution of words according to their size was alike for both patterns in this position. It is worth noticing how strikingly similar the overall picture is to that of Study 1 (Figure 3.2).

To test the observation statistically, I used a GLM analysis again, with factors position, word size, speaker, pattern type, and with a 3-way interaction position*word size*pattern type. As in Study 1, the main interest was the 3-way interaction. Surprisingly, this time the interaction turned out highly significant; [$F(21,681) = 9.99$, $p < 0.001$]. The discrepancy between the statistical results in Studies 1 and 2 may be partly due to the much larger number of observations in Study 2. The conclusion is then that the distribution of words according to their size with respect to their position within the phrase is likely to contribute to some extent to the difference in articulation rate variation in patterns ABC and BAC. This is obviously an important issue. To be able to estimate to what extent the distribution of words of different size within the phrase contributes to the difference between patterns ABC and BAC, further statistical analysis would be necessary. However, such analysis is beyond the scope of this thesis.

Study 2 also demonstrated that word class is an important factor affecting the articulation rate of individual phonological words, particularly, if they contain only content or only function words. This may be relevant for the present investigation of ABC and BAC patterns. A general tendency found in my data was that phonological words containing only function words were significantly faster than those containing only content words. My hypothesis is then that in pattern ABC, more words in position 1 will be from the function word category than in pattern BAC, and in pattern BAC, more words from the function word category will be in position 2 in comparison with pattern ABC. Figure 6.10 shows the distribution.

Figure 6.10 **Distribution of words according to their word class category for different positions within the intonation phrase.**



It is clear from the graph that the first part of my hypothesis is false; in position 1, the distribution of words in all three word class categories is very similar in patterns ABC and BAC. The second part of the hypothesis, however, seems to hold; in position 2, pattern BAC has twice as many words in the category of function words as pattern ABC (40% as opposed to 20%). Moreover, the content word category (i.e., the slowest category) is somewhat more strongly represented in position 2 in the ABC pattern than in the BAC pattern (60% as opposed to 50%). Both these facts may have contributed to the fact that the second word in pattern BAC is the fastest within the intonation phrase. However, it does not account for the first word being the fastest in the ABC pattern.

A GLM analysis, with factors position, speaker, word class, pattern type, and a 3-way interaction position*word class*pattern type, was designed to find out whether the discrepancies between the two patterns observed above are statistically significant. Again, the 3-way interaction was of main interest and it was highly significant; [F (12,693) = 15.55, $p < 0.001$]. It is likely that this significant result reflects mainly the difference between the second positions in

patterns ABC and BAC. If this is so, then the articulation rate variation in pattern BAC may be to some extent determined by the distribution of function words within the intonation phrase.

To summarize the findings on patterns ABC and BAC, the analysis suggested that, overall, the two patterns were comparably strong; they had all positions significantly different from each other in their articulation rates. The examination of two aspects of the linguistic structure of the phrases, the size of component words in syllables and the word class they belong to, did not lead to any conclusive results. Although there are some indications that they both affected the difference in articulation rate variation in the two patterns, it is not clear to what extent. Further statistical analysis would be necessary to resolve this issue. Thus, the hypothesis developed in Study 1 that the ABC pattern is the default is weakened; at this stage, it is not possible to determine whether either pattern is the default pattern.

Finally, as in Study 1, I examined other non-rallentando patterns than BAC in 3-word phrases (Table 5.4), and non-rallentando patterns in 4-word phrases (Table 5.5). I considered mainly the linguistic and discourse structures and pausing. In Table 5.4, pattern ACB stands out as the most common pattern after ABC and BAC in all three tasks. In the reading task there was an intonation phrase within which this pattern occurred in all four speakers. It turned out that the word in the second position (B) was *skutečně* ('really') and it seemed to be slightly emphasized. I found two other similar cases in the material, one involving the word *zřejmě* ('apparently'), the other involving the word *spíš* ('rather'). Since it is known that emphasis makes the duration of words longer, it is plausible that this made the intonation phrase surface with the ACB pattern. However, most of the cases with pattern ACB have a much simpler explanation – the absolute values for the articulation rate of B and C were very similar (the difference between the two did not exceed 0.3 syll/s in most cases) and this made them, in fact, close to the rallentando pattern ABC.

The other relative common non-rallentando pattern was BCA (Table 5.4). However, no regularities which could explain this pattern became apparent. The same was true about 4-word non-rallentando patterns. To conclude, the analysis suggested that there might be some further factors apart from word size and word class which could contribute to articulation rate variation, in particular, the discourse structure.

6.3 CONCLUSIONS

The conclusions in terms of testing the hypotheses introduced in section 4.1 are as follows:

Hypothesis 2.2:
Articulation rate variation within the intonation phrase reflects a regular pattern of rallentando (progressive slowing down) in all four speakers and all three tasks.

Hypothesis 2.2 was confirmed; in Czech a slowing down seems to be the overall pattern of articulation rate variation within the intonation phrase. This result applied to all four speakers, and it was consistent across reading, paraphrasing a story and spontaneous speech.

Hypothesis 3.2:
The articulation rates of words in all positions within the intonation phrase are significantly different from each other in all four speakers and all three tasks.

When the data from all speakers were combined together, the slowing down was non-linear, apart from intonation phrases containing only two words. Hypothesis 3.2 was confirmed for the dataset as a whole; all the positions were significantly

different from each other, with the mean articulation rate of the last word within the phrase being noticeably slower than all the previous words.

The part of the Hypothesis 3.2 referring to individual speakers was tested only for 3-word phrases, since only in this group of phrases did the statistics indicate possible differences among speakers. Different degrees of non-linear slowing down were found. In two speakers all the words in 3-word phrases had significantly different mean articulation rates, while for the other two speakers the mean articulation rates for the first two words were not statistically different. The tendency of the last word to be significantly slower than all the previous words was consistent.

Hypothesis 3.2 was not studied for individual tasks since the statistical analysis suggested that no significant differences exist between them as to what kinds of patterns of rate variation are typical for them.

Hypothesis 4.2:
The size of the word, counted in syllables, significantly affects the word's articulation rate in all four speakers and all three tasks.

Hypothesis 4.2 was confirmed. As the number of syllables in the word increases, its articulation rate increases as well. This tendency was consistent across speakers and all three tasks. However, the trend was non-linear. The mean articulation rates of words containing three or more syllables were not statistically different from each other. Monosyllabic and disyllabic words were the most distinct categories.

Hypothesis 5.2:
The word class significantly affects the word's articulation rate in all four speakers and all three tasks.

The factor of word class also significantly affected the articulation rate of words in all four speakers and all tasks. The category of function words was significantly faster in mean articulation rate than the category of content words, which were the slowest of the three categories. The remaining category of content+function word(s) was significantly faster than content words and generally seemed slower than the function word category. Since the difference was significant only in 3-word phrases, no strong conclusions can be drawn about the relationship between these two categories.

Position of the word within the intonation phrase and its word class category might have to some extent a confounding effect on the articulation rate of words. Function words, which were shown to be the faster word category, tend to be in the first position of the phrase and content words become more frequent towards the end of the phrase. At the same time, words in the first position tend to be the fastest of all the words in the phrase and the final words the slowest. Further statistical analysis would be necessary to resolve to what extent the effects of position and word class are confounded.

Hypothesis 6.2:
Speakers differ significantly from each other in their articulation rate of words.

Although the factor of speaker proved significant for the articulation rate of words, Hypothesis 6.2 was not fully confirmed. Not *all* the speakers contributed to the overall significant result; there were only two speakers whose mean articulation rates were statistically different in a consistent way.

Hypothesis 7.2:
Word articulation rate is significantly affected by the kind of task the speaker is carrying out. Reading is significantly faster than spontaneous speech, and spontaneous speech is faster than paraphrasing.

Here the results are not clear. The examination of means suggested that in all three groups of phrases, the fastest articulation rate occurred in spontaneous speech, not in reading. However, the factor of task was significant only in 3- and 4-word phrases. In 3-word phrases, this was due to spontaneous speech being significantly faster than paraphrasing, and in 4-word phrases, due to spontaneous speech being significantly faster than reading.

The analysis of the relative strength of the factors suggested that articulation rate of words is most strongly affected by their position within the intonation phrase and the number of syllables they contain, and less so by word class. Speaker and, particularly, the kind of task seem have the weakest effects.

CHAPTER 7

MATHEMATICAL MODELLING OF ARTICULATION RATE

7.1 INTRODUCTION

The aim of this chapter is to provide a quantitative model of articulation rate, based on the data in Study 2. Such a model can be used for estimating and, subsequently, predicting articulation rate. In this chapter, I shall concentrate on the stage of estimating articulation rate.

Estimating the dependent variable involves establishing the relevant explanatory variables and the relationships between them. The first part of this procedure has been largely achieved in Chapter 6. In this chapter, the focus is on a more detailed examination of how the explanatory variables relate to each other, so that a model for articulation rate that would best fit the observed values of articulation rate can be designed.

The type of model chosen was the sums-of-products model, which is a case of a class called additive-multiplicative models (van Santen, 1992). It is a linear model with multiple explanatory variables. A general form of such a model with say, three explanatory variables, or factors, can be represented by the following equation:

$$Y = a + b X_1 + c X_2 + d X_3 + \varepsilon$$

where Y is the dependent variable, X_1, X_2, and X_3 are explanatory variables, a, b, c and d are coefficients, or parameters, and ε is the error term. The coefficients are

constants estimated from the data set by the method of least squares in multiple regression analysis (e.g., Ferguson & Takane, 1989). The error term is needed in the equation, if the aim is to explain the dependent variable completely. It contains both variation that would be accounted for if some additional factors were included in the model, and random error. From the point of view of the model geometry, the regression gives the best-fitting hyperplane.

Assumptions for the multiple (linear) regression analysis are: (i) all the coefficients, or parameters, appear linearly[1]; (ii) the dependent variable is continuous and for each value of the independent variable, its distribution is normal; (iii) the variation around the regression hyperplane is the same for all values of each X (constant or homogeneous variance); (iv) successive observations of the dependent variable are uncorrelated; and (v) the independent variables are not highly correlated.

For the data in Study 2, assumptions (iv) and (v) were fulfilled (see the discussion in section 6.1). Homogeneity of variance and normal distribution were checked while designing the model itself and used as one of the criteria for accepting/rejecting the model.

In the present study, the dependent variable was articulation rate, and the explanatory variables were mainly those which in the earlier GLM analysis of Study 2 data (Table 6.1, p.124) proved to significantly affect the articulation rate of phonological words, i.e., word size in syllables ('sylls'), its position ('pos') and the word class ('wclass'). Factors 'speaker' and 'task' were not included as explanatory variables in the model for the following reasons: (i) in comparison with other main explanatory variables, they affected articulation rate to a much lesser degree, and, in fact, 'task' was not statistically significant at all in 2-word phrases (see Table 6.1); (ii) in comparison with 'position', 'size' and 'wclass',

[1] In this, I follow the standard approach (e.g., Ratkowsky, 1990). Some statistical handbooks reserve the term linear regression for models in which the relationship between the independent and the dependent variables is linear (e.g., Mason *et al.*, 1983).

factors 'speaker' and 'task' are factors of a different kind; while the first three are linguistic variables and as such they are the main concern in this thesis, the latter two are extra-linguistic variables. Moreover, the factor 'speaker' cannot really be altered by anyone; in other words, each speaker has his own characteristic overall articulation rate, which is, under comparable circumstances, relatively stable.

In all the statistical analysis described in Chapters 3 and 6, I treated groups of 2-word, 3-word and 4-word phrases separately, given that the coding for position is not directly comparable across these three groups of phrases (see the discussion in section 2.7). In modelling articulation rate, however, it is preferable to design a single model that is based on all the data as a whole, rather than having several (in this case three) separate models, since such a model is more general. In order to compensate for the differences in status of the levels within the variable of position, an additional variable of the number of words in the phrase ('nwords') was included in the model. Further justification for using 'nwords' was that when a GLM model was run with this variable added to all the main factors used before, this factor proved significant; thus it seems that for articulation rate of a word, it does matter how many words there were in the phrase it came from; [F (2; 2,101) = 114.5, $p < 0.001$]. However, a post-hoc Tukey test revealed that not all three groups of phrases were significantly different from each other; the mean articulation rate for 2-word phrases was statistically different from the means for both 3-word and 4-word phrases; the latter two were not statistically different.

Six mathematical models of articulation rate are considered in this chapter. The details about each of them and the reasons for the choice of these particular models will be discussed in the course of this chapter. Computations were conducted mainly by using the General Linear Model procedure (General Factorial) and Linear Regression, both in SPSS, version 7.5.1. In the GLM, 'pos' 'sylls', and 'nwords' were treated as covariates, (i.e., numerical variables);

'wclass' was treated as a factor since it is a categorical (nominal) variable.[2] The full results are included in Appendix 10.

The GLM analysis provided estimates for regression coefficients and also coefficients of determination (R^2), which provided a measure of success of the model by giving a percentage of variation accounted for by the explanatory variables used in the model. Both of these measures can be seen as measures of goodness of fit. However, for selecting the best model, adjusted R^2 was used since this measure, unlike R^2, compensates for the number of parameters in the model and thus makes comparison of models with a different number of parameters possible. A χ^2 test was used to find out whether the improvement/decline in success of two models was significant (the calculation for this test involved simply deducting the sums of squares for the error and for the degrees of freedom in the two models in question, and then finding the significance level in a χ^2 table). However, in some cases it was not possible to provide the significance level since the number of degrees of freedom in the two models was the same (so the difference was equal to zero). In the comparison of individual models, the values for standard error of estimate and the minimum and maximum residuals were also taken into account. Finally, histograms and scatter plots were used to check the normality of the distribution of residuals (which approximate the error term) and the homogeneity of variance in the dependent variable. I found that in all six models the variance was unequal around the regression line; the greater the values for articulation rate, the greater the dispersion around the regression line. Since this goes against the assumptions for regression analysis, an attempt was made to correct this and to equalize the variation along the line by using a logarithmic transformation (natural logarithm) of the predicted values for articulation rate. This logarithmic transformation was conducted for all six models.

[2] Recall from Chapters 2 and 4 that in the previous statistical analyses, all the variables were treated as factors, i.e., including position and size (syllables), since post-hoc tests for significance of means would not be possible if they were treated as covariates. However, in the present analysis, they had to be coded as numerical since the aim was to design a numerical model.

The error term (ε) reflects unexplained variation in observed articulation rate. Modelling the error term might improve goodness of fit in the long term. However, this was not carried out at this stage, since the aim was just to find the simplest model that would explain as much variation as possible, given several selected explanatory variables.

As I mentioned above, one of the explanatory variables, the word class, was a nominal variable. The fact that it involved three levels created a problem, since having a nominal variable with more than two levels prevents it from being included in a simple regression analysis as it stands. There are two basic ways of solving this problem. The first one involves recoding the original, three-level, nominal variable into two binary dummy variables; the second, a more complicated way, involves designing a model where a single intercept is replaced by three intercepts, each for one level of the original nominal variable; interaction terms between these levels, now variables themselves, and the remaining explanatory variables are included as well. I used both options. The models will be explained in more detail in the following sections.

7.2 MODEL 1

The first model involved recoding the nominal variable of word class into two dummy variables.

Recall from the description of the GLM analysis (Chapter 4, section 4.7.3) that the three levels within this factor were:

1	C	content word(s) only
2	C + F	content word + function word(s)
3	F	function word(s) only

This coding was needed at that stage, since phonological words did not always contain purely content or purely function words; often they were a mixture of both. The three-way coding allowed me to see whether this mixed category is distinct with respect to articulation rate, or whether it behaves similarly to one of the other two categories (i.e., whether it is not statistically different from them in mean articulation rate). [3] The analysis showed that in two of the three groups of phrases, this mixed category behaved similarly to the function word category (see section 6.2.2.3).

Dummy coding involved assignment of the weights 1 and 0 to represent membership in the individual levels of the factor. The original coding became represented by the combination of two dummy variables, which were called 'cont' and 'func', in the following way:

		'cont'	'func'
1	C	1	0
2	C + F	1	1
3	F	0	1

The model using these two dummy variables was then the simplest linear model, given all the selected explanatory variables. It is represented in the equation below:

$$AR = a + b\,sylls + c\,pos + d\,cont + e\,func + f\,nwords + \varepsilon$$

However, as I mentioned in section 7.1, a logarithmic transformation of the observed articulation rate was performed before fitting the model, since this helped to correct for the inhomogeneity of variance found when the predicted values for articulation rate were plotted against the observed articulation rate. A

[3] The numerical coding of the three levels of word class is, of course, completely arbitrary.

modified linear regression equation is then (subscript 1 indicates that this equation is for Model 1):

$$\log (AR_1) = a + b\,sylls + c\,pos + d\,cont + e\,func + f\,nwords + \varepsilon$$

The model for estimated articulation rate (AR') derived from the data can then be expressed as:

$$\log (AR_1') = 1.23 + 0.16\,sylls - 0.17\,pos + 0.21\,cont - 0.002\,func + 0.14\,nwords$$

The adjusted R^2 was 0.45, the standard error of estimate was 0.253; the minimum residual was −1.14 and the maximum residual 0.91.[4] t tests showed that all coefficients were significantly different from zero ($p < 0.0001$), apart from the coefficient for factor 'func'. Normally this would justify excluding the variable from the equation. However, since variable 'func' represents both levels C+F and F of the word class, not just a single variable, it should stay in the model.

Non-significance of the coefficient for 'func' suggests that distinguishing three levels within word class might not be of any advantage in modelling articulation rate. Thus a new model was designed (Model 2), where word class was converted into a binary category.

[4] The values for standard error of estimate and for the minimum and maximum residuals are also in a logarithmic scale. Since the main purpose for using these values here was the comparison of different models and they all involved a logarithmic scale, converting the values back into the scale of the number of syllables per second (by using the exponential function) was not necessary at this stage.

7.3 MODEL 2

In Model 1, variable 'func' treated level 2 (C+F) and level 3 (F) from the original classification as the same. A similar approach was adopted in designing Model 2. The original word class classification was recoded into a binary classification ('wclassbin'); the original level C was one separate level, the original levels C+F and F together were the other separate level. (Recall that only in 3-word phrases were levels 2 and 3 significantly different in the GLM analysis; section 6.2.2.3.). The GLM analysis gave the following equation:

$$\log (AR_2') = 1.32 + 0.13\, sylls - 0.17\, pos + 0.1\ wclassbin + 0.14\, nwords$$

When comparing the success of this model to Model 1, the present model was found less successful. The standard error of estimate was 0.261 (as opposed to 0.253), the minimum residual was −1.18, the maximum residual was 0.89, and the adjusted R^2 in Model 2 was 0.42 (as opposed to 0.45 in Model 1). A χ^2 test showed that this is a statistically significant drop in success ($p < 0.01$). Thus it can be concluded that treating the factor of word class as binary did not improve the model of articulation rate, rather the contrary.

7.4 MODEL 3

In the search for the best model, the possibility of adding two-way interactions to the main factors in the model was briefly considered. The model did improve the goodness of fit to the observed articulation rate; however, the improvement was only very slight, at the expense of a considerably increased complexity of the model. Accepting this model would go against the criterion of parsimony, which should be kept in mind in any sort of modelling. The failure of the model to achieve greater improvement was not surprising, given that none of the two-way

interactions involving the factors used in this model were significant in the previous GLM analysis (section 6.2.1).

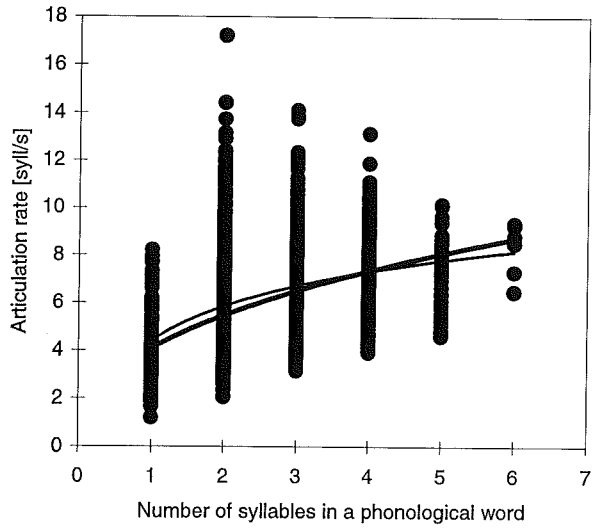
The next step in modelling articulation rate involved looking more closely into the nature of the relationship between articulation rate and the individual explanatory variables. As I showed in sections 6.2.2.1 and 6.2.2.2, articulation rate was non-linear with respect to two factors, position (for 3-word and 4-word phrases), and word size in syllables (for all three groups of intonation phrases). Given the fact that all three groups of phrases were combined for modelling articulation rate and that the original coding of position was preserved, articulation rate and position were in almost a linear relationship.[5] Therefore, at this stage, I shall discuss only the factor of word size ('sylls'), as the only candidate for transformation or linearization.

In order to find out what kind of transformation was needed, the curve estimation procedure in SPSS was used. The choice of the best fitting model was based on comparing the coefficients of determination, R^2. However, linguistic plausibility of various models was also taken into account. Significance levels for the differences between the models could not be obtained, since the degrees of freedom between the models were identical.

A word's articulation rate tended to increase as the number of syllables in the word increased. In Chapter 4, I pointed out that this tendency was non-linear in a sense that the mean articulation rates of words containing three or more syllables were progressively less different from each other. Figure 7.1 shows the plot for all three groups of intonation phrases combined.

[5] The best fit was, in fact, a power function but the improvement in the coefficient of determination (R^2) was only marginal in comparison with the linear function (17.9% and 16.7% respectively).

Figure 7.1 **Scatterplot of the articulation rate as a function of size in syllables. The thick line is the curve for the power function AR' = 4.09 (sylls)$^{0.42}$; the thin line is the curve for the logarithmic function AR' = 4.4 + 2 log (sylls) .**



As I mentioned in section 7.1, a linear relationship between the dependent and independent variables is assumed in linear regression analysis, but the analysis is reasonably robust to a certain degree of violation of its assumptions. In any case, linearization of the variable 'sylls' might improve the success of the model. The data for 'sylls' in Figure 7.1 suggest two functions, namely, logarithmic (thin curve) and power (thick curve). The curve estimation procedure in SPSS suggested that a power function gives a better match for the data on articulation rate than a logarithmic function (R^2 = 0.23 for a power function; R^2 = 0.16 for a logarithmic function; for comparison, for a linear function, R^2 = 0.13). The equation expressing the articulation rate of a word as a power function of word size is as follows:

$$AR' = 4.09\ (sylls)^{0.42}$$

Thus, prior to running the next model, Model 3, a transformation of the variable 'sylls' was performed by raising it to the power 0.42. The regression analysis then gave the following equation:

$$\log (AR_3') = 0.61 + 0.71\ (sylls)^{0.42} - 0.16\ pos + 0.22\ cont - 0.003\ func + 0.13\ nwords$$

Evaluation of this model versus Model 1 suggests this model to be better than Model 1 with respect to both the standard error of estimate (Model 3: 0.247), the minimum residual (–1.07), the maximum residual (0.90), and the adjusted R^2 (0.48). The significance level for this improvement could not be obtained, since, again, the difference in the degrees of freedom between the two model was equal to zero.

7.5 MODEL 4

Although the curve estimation procedure suggested that the logarithmic function for the variable 'sylls' was not as good in terms of the goodness of fit as the power function, a logarithmic transformation of 'sylls' was also performed since, from the linguistic point of view, a logarithmic function is also a plausible function; it grows more slowly as x values increase. This is exactly what happens with articulation rate as the number of syllables in the word increases. In words with more than about three or four syllables, articulation rate 'flattens out' because of the time constraints on articulatory movements (e.g., Kohler, 1991). From this perspective, a logarithmic function is in fact a more plausible function than the power function with the power of 0.42, since the latter increases more steeply for higher values of x.

Thus in the new model, Model 4, a logarithmic transformation of the variable 'sylls' (natural logarithm) was applied before the regression analysis. The equation based on this analysis is as follows:

$$\log (AR_4') = 1.27 + 0.43 \log (sylls) - 0.16\ pos + 0.22\ cont - 0.002\ func + 0.13\ nwords$$

As to the goodness of fit to the observed articulation rate, Model 4 turned out, surprisingly, to be better than Model 3. In fact, it was found to be the best model of all four models used in the analysis so far. The standard error of estimate was 0.242, the minimum residual –1.02, the maximum residual 0.90, and the adjusted R^2 was 0.50. Again, since the difference in the degrees of freedom between Model 3 and Model 4 was equal to zero, a significance level cannot be provided. The residuals were normally distributed.

7.6 MODEL 5

So far, modelling articulation rate involved dummy variables 'cont' and 'func', and Models 1 - 4 represented attempts to refine the modelling so as to explain more of the variance in the dependent variable, while keeping the model as simple as possible.

As I explained in section 7.1, the other way of solving the problem with a nominal variable with more than two levels, is to use a more complicated model, where a single intercept is replaced by three intercepts, each for one level of the original nominal variable, and where two-way interactions between these levels, now variables themselves, and the remaining explanatory variables are included as well.

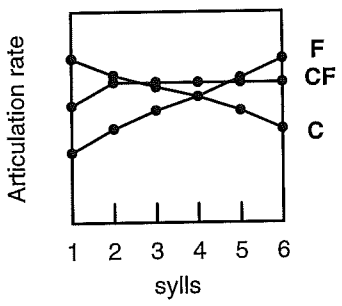
This multi-dimentional model used variable C as the reference level. It has the following form (assuming a prior logarithmic transformation of the observed articulation rate that had to be performed):

$$\log(AR_5) = a*C + b*CF + c*F +$$
$$+ d*C*sylls + e*C*pos + f*C*nwords +$$
$$+ g*(C–CF)*sylls + h*(C–CF)*pos + i*(C–CF)*nwords +$$
$$+ j*(C–F)*sylls + k*(C–F)*pos + l*(C–F)*nwords + \varepsilon$$

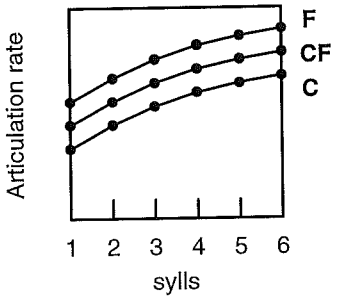
The first line represents three separate intercepts, one for each level of the original word class category. The second line stands for slopes of variable C with respect to all the remaining variables, i.e., in the plane for 'sylls', in the plane for 'pos' and in the plane for 'nwords'. The following, third, line represents differences in slopes between variable C and variable CF, with respect to the remaining variables in the model. Finally, the last line stands for differences in slopes between C and F, again with respect to the remaining variables.

This model, unlike Models 1 - 4, allows variation in articulation rate to differ across categories C, CF and F in all three dimensions, i.e., 'sylls', 'pos' and 'nwords'. This possibility is shown in the schematic representation (a) below (showing only one plane, articulation rate vs. 'sylls'); the absence of such differences in variation is shown in (b):

(a)



(b)



In this model, Model 5, 'sylls' were not transformed in any way, so that comparison would be possible between this model and Model 1, which also used untransformed 'sylls' but was simpler than Model 5 in the sense that no

interactions were involved (i.e., the possibility (b) was assumed there). The model that the GLM analysis arrived at is expressed by the following equation:

$$\begin{aligned}\log(AR_5') = {} & 1.12 * C + 1.68 * CF + 1.47 * F + \\ & + 0.18 * C * sylls - 0.15 * C * pos + 0.1 * C * nwords + \\ & + 0.19 * CF * sylls - 0.21 * CF * pos + 0.19 * CF * nwords + \\ & + 0.067 * F * sylls - 0.1 * F * pos + 0.05 * F * nwords\end{aligned}$$

When comparing Model 1 and Model 5, Model 5 seems more successful than Model 1 in all parameters used for the comparison. (Model 5: the standard error of estimate = 0.237; the minimum residual = –0.96; the maximum residual = 0.88; the adjusted R^2 = 0.47. Model 1: the standard error of estimate = 0.253; the minimum residual = –1.14; the maximum residual = 0.91; the adjusted R^2 = 0.45.) However, a χ^2 test showed that the improvement in adjusted R^2 was not statistically significant ($p > 0.05$).

7.7 MODEL 6

Given the result of the χ^2 test on Models 1 and 5, it is not surprising that Model 5 was found less successful than Model 4, the model which was most successful so far. Recall that Model 4 was an improvement to Model 1 by including a logarithmic transformation of 'sylls'. To be able to make a fair comparison with Model 4, a logarithmic transformation of 'sylls' was applied to Model 5 as well. The resulting model became Model 6. With the coefficients included, the equation has the following form:

$$\begin{aligned}\log(AR_6') = {} & 1.19 * C + 1.65 * CF + 1.56 * F + \\ & + 0.41 * C * \log(sylls) - 0.14 * C * pos + 0.1 * C * nwords + \\ & + 0.48 * CF * \log(sylls) - 0.08 * CF * pos + 0.19 * CF * nwords + \\ & + 0.22 * F * \log(sylls) - 0.1 * F * pos + 0.05 * F * nwords\end{aligned}$$

The standard error of estimate for Model 6 was 0.237, the minimum residual –0.96, the maximum residual 0.88, and the adjusted R^2 was 0.51. In all these parameters, Model 6 was more successful than the previous model, Model 5. This was expected since we already know that a logarithmic transformation of 'sylls' improves the model.

However, Model 6 also seemed to do better than Model 4, which did involve a logarithmic transformation of 'sylls'. But once again, a χ^2 test demonstrated that the improvement was too small to be statistically significant ($p > 0.05$).

7.8 CONCLUSION

Table 7.1 below summarizes the results for a number of parameters used for the estimation of goodness of fit and for the comparison of the individual models.

Table 7.1 Estimates for goodness of fit in Models 1-6

	Standard error	Adj. R^2	Minimum residual	Maximum residual	No. of parameters in the model (and *df*)
Model 1	0.253	44.8%	-1.14	0.91	5 (5)
Model 2	0.261	41.6%	-1.18	0.89	4 (4)
Model 3	0.247	47.8%	-1.07	0.90	5 (5)
Model 4	**0.242**	**49.6%**	**-1.02**	**0.90**	**5 (5)**
Model 5	0.247	47.3%	-0.96	0.88	7 (12)
Model 6	0.237	51.4%	-0.96	0.88	7 (12)

N.B. The residuals were normally distributed in all six models.

The choice of the best model is normally driven by two main criteria: (i) the success of the model in terms of the amount of variance in the dependent variable accounted for by the explanatory variables, and (ii) the criterion of parsimony.

Depending on research aims, one can be more important than the other, but, ideally, both should be fulfilled. From this perspective, the best model for the articulation rate observed in Study 2 is Model 4 (in Table 7.1 in bold), i.e., the model in which five explanatory variables were used, two of them being dummy variables for the original word class, and in which both 'syll' and the observed articulation rate were transformed logarithmically. For convenience, I shall repeat the relevant equation below:

$$\log (AR_4') = 1.27 + 0.43 \log (sylls) - 0.16\ pos + 0.22\ cont - 0.002\ func + 0.13\ nwords$$

This model accounts for 50% of variance in the observed articulation rate. For this final model, the inverse function to the natural logarithm (i.e., the exponential function) was applied in order to obtain the minimum and maximum residuals in the original scale of the number of syllables per second. For the minimum residual the result was 4.85 syll/s; for the maximum residual the value was 9.09 syll/s.

Adding complexity to the equation in terms of interactions between each level of the word class and the remaining explanatory variables (Models 5 and 6) did not improve the goodness of fit to the observed articulation rate significantly.

The model suggested in this chapter could be used for predicting articulation rate for, say, another speaker. However, several more steps would have to be added in the analysis, namely, standard error of prediction, confidence intervals and converting the predicted articulation rate back to the scale of syllables per second by using the exponential function. It was not my intention to achieve this in this thesis.

CHAPTER 8

SUMMARY AND DISCUSSION

8.1 INTONATION PHRASE AS A DOMAIN OF ARTICULATION RATE VARIATION

The results reported in Chapters 3, 5 and 6 suggest that articulation rate does have, at least in Czech, a specific domain of variation, and that this domain is the intonation phrase. The term 'domain of variation' was described in the first chapter by two possible scenarios; either (i) articulation rate is constant within the domain and changes in rate happen *at the unit boundaries*, or (ii) a regular pattern of rate change happens *within* the domain. The first case clearly does not apply to Czech. Articulation rate within the intonation phrase was not constant in my study; it varied, typically between 20% and 40% from the mean.[1] If proved for languages in general, this would contradict the tentative assumption noted by Roach (1982) that tempo variation within the tone unit (i.e., the intonation phrase) is exceptional. He says:

> "It is assumed that the tempo changes will usually be manifested in terms of differences between the tempo of one tone unit and another, and will not usually be found within the tone unit (though since Crystal and Quirk, 1964, have pointed out the phenomena of accelerando and rallentando we must accept that this does sometimes happen)". (Roach, 1982:76)

[1] As I pointed out in section 3.2, there does not seem to be an established criterion for what should be regarded as 'small' or 'considerable' variation in articulation rate. However, most studies consider a variation over 20% as considerable. If a perceptual threshold of just noticeable differences in articulation rate is used for evaluation, the variation in my study is well beyond this threshold.

The statistical analysis presented in this thesis showed that in Czech, not only is the variation in articulation rate not bound exclusively to the intonation phrase boundaries, but it is not random. Rather, it tends to follow a systematic pattern of a progressive slowing down (rallentando) within the intonation phrase.[2]

A further persuasive piece of evidence for non-randomness of articulation rate variation was provided by the comparison of the rate of individual phonological words across different speakers reading the same text. The degree of similarity between the speakers was striking.

The pattern of slowing down was consistent across all speakers in two different studies, and it also seems to be common to different speech tasks, at least to reading a story, paraphrasing a story and spontaneous speech. However, the slowing down was non-linear; it was rather gradual across non-final words, with the last word within the phrase being noticeably slower than any other preceding words within the same phrase. Yet, when a large database was analysed, all words in three-word phrases, i.e., including non-final words, were statistically different from each other in their mean articulation rates. In four-word phrases, the evidence for non-final words being significantly different from each other was weaker; apart from the final word being significantly slower than all the non-final words, only the first and third words were statistically different from each other.

A strong tendency for the last word being the slowest within the intonation phrase suggests that the pattern of articulation rate variation is partially determined by phrase-final lengthening. To turn it the other way round, final lengthening in Czech seems to be a property of intonation phrases. This is in accord with the finding of Beckman and Edwards (1990) for English and supports the view that the constituents which trigger final lengthening are phonological in nature

[2] The terms rallentando and accelerando as used in this thesis are not synonymous to rallentando and accelerando as defined in Crystal's complex tempo classification (1969), based on perceptual analysis.

(Gussenhoven & Rietveld, 1992). This relatively recent view contrasts with a more traditional and widespread claim that final lengthening is governed by syntactic structure. Although my study does not support this claim, I considered the possibility that the clause and/or interpause stretch might be a hierarchically higher-level domain(s) of phrase-final lengthening (in most cases, pauses and clause boundaries were aligned with intonation-phrase boundaries). The results indicated that the clause is not a domain of final lengthening; for the interpause stretch, the results were inconclusive. It does seem that in Czech, phrase-final lengthening tends to be greater in the last word of the final intonation phrase within the interpause stretch than in the last words of non-final intonation phrases. This provides some justification for the term 'prepausal lengthening' which is sometimes used as a synonym to phrase-final lengthening. More data is necessary to make a stronger claim about phrase-final lengthening within the interpause stretch.

Regarding clauses, it seems likely that the reason why the phenomenon of final lengthening has been commonly attributed to the syntactic boundaries (e.g., Klatt, 1975; Cooper & Paccia-Cooper, 1980; Selkirk, 1984; Crystal & House, 1988) is that syntactic phrases are normally coextensive with intonation phrases (e.g., Fletcher, 1991), an observation which certainly applied also to my data.

In principle, there is no reason why articulation rate should have only one domain of variation. In this thesis, I considered two other possible domains: the clause and the interpause stretch. I have already mentioned that, in most cases, pauses and clause boundaries were aligned with intonation-phrase boundaries. Those clauses and interpause stretches which coincided with intonation phrases had, of course, the same pattern as the intonation phrase. On those clauses and interpause stretches that contained more than one intonation phrase, I tested the hypothesis that the interpause stretch and/or the clause might be a hierarchically higher domain(s) of variation in articulation rate. However, no consistent pattern was found in either of these units.

The finding that articulation rate may be organized at the level of intonation phrases makes a lot of sense from both the linguistic and speech production points of view. The intonation phrase is traditionally considered a domain of the intonation contour. However, some researchers claim explicitly that the intonation phrase is a rhythmical unit as well. For instance, Palková (1994:288) suggests that the intonation phrase is:

> "probably hierarchically the highest unit to which rhythmical properties of speech apply"

and that in speech production, this manifests itself through constraints on the possible variation in intonation phrases both with respect to the number of syllables and the number of phonological words they contain. Selkirk (1984) considers the intonation phrase a domain with respect to which patterns of rhythmic prominence are defined. This thesis shows that the rhythmical properties of intonation phrases also involve regular variations in articulation rate, and thus supports the view that the intonation phrase is one of the fundamental units in the temporal patterning of speech (e.g., Butcher, 1981) and a psychologically real unit of speech production.

Some researchers have suggested that the intonation phrase (tone unit) plays an important role in speech planning. An initial formulation of the tone unit in these terms was put forward by Boomer (1965) in his study of hesitation pauses (his 'phonemic clause'). Boomer proposed the phonemic clause to be the encoding unit of speech at the grammatical level, as both filled and unfilled pauses occur most frequently at the beginning of the phonemic clause. However, O'Connell and Kowal (1983) point out a serious flaw in Boomer's approach; although his phonemic clause was theoretically defined phonologically by a single primary stress and terminal juncture, it was, in fact, determined circularly from the unfilled pauses in the recordings.

Claims about the importance of the intonation phrase as a planning unit were made also by, for instance, Halliday (1967), Lehiste (1970) and Hirst (1991). Additional support comes from the research on slips of the tongue by Boomer and Laver (1968/1973), who found that the target and the origin of a tongue-slip are both located in the same tone group and that the typical slip involves interference from the tonic word *before* it is uttered. According to Boomer and Laver,

> "the tone-group is handled in the central nervous system as a unitary behavioural act, and the neural correlates of the separate elements are assembled and partially activated, or 'primed', before the performance of the utterance begins" (Boomer & Laver, 1968/1973:127).

However, it is now understood (e.g., Levelt, 1989) that there is no single planning unit but rather that different levels of speech production use different processing units. The intonation phrase is a high-level planning unit used for planning an intonation structure, which then undergoes phonetic interpretation. At the stage of phonetic interpretation an F_0 contour is generated. A related view is presented, for instance, in Selkirk (1984), in her discussion of how prosodic structure relates to other aspects of linguistic structure, in particular syntactic and semantic.

Selkirk (1984) distinguishes three levels of mapping from syntax to phonology: (i) surface syntactic structure, (ii) intonated surface structure, and (iii) intonated and rhythmated surface structure. She proposed that intonated surface structure designates the intonational phrasing of the sentence and tonal contours (represented autosegmentally) and the assignment of certain of these tonal elements to particular constituents of surface structure. In other words, this level entails a more abstract, phonological representation of intonation, in terms of sequences of tonal elements, such as pitch accents and boundary tones. Such a representation is the basis of the autosegmental theory of intonation (Pierrehumbert, 1981; Beckman & Pierrehumbert, 1986). The intonated and rhythmated surface structure involves also a metrical grid. The resulting abstract representation, i.e., underlying phonological representation, is then mapped onto a

Oxford University
Phonetics

'concrete', low level phonetic representation. In Pierrehumbert's theory this involves phonetic implementation, in which pitch accents are targets of the movements in fundamental frequency.

The finding of a regular patterning of articulation rate reported in this thesis suggests that the intonation phrase, in addition to its role in a high-level planning of the intonation contour, is further involved in the generation of temporal structure, in particular, in structuring articulation rate variation. A progressive slowing down in articulation rate within the intonation phrase seems to be a general underlying pattern of articulation rate variation, at least in Czech. By 'general' I mean that it applies both across different speakers and across different tasks. By 'underlying' I mean that it is likely to be an unintentional, inherent property of speech on which intentional changes in rate can be built. It must be remembered that the pattern is a *tendency*, so on the surface it may emerge with a relatively large variation.

An interesting parallel emerges between rallentando in articulation rate and declination in the sense that both involve a progressive decline, and both seem to have the intonation phrase as their domain.[3] A question arises whether they are triggered by similar phenomena, and, related to that, whether they might be planned at the same stage in speech production. A still common view of declination, a phenomenon attested in many languages (Vaissière, 1983), is that it is an involuntary, or automatic process, caused by certain physiological aspects of speech production (Lieberman, 1967; Ohala & Ewan, 1973; Collier, 1975, 1983; Maeda, 1976). This would suggest that declination is generated at a low level in speech production. However, more recently, there has been increasing support for the suggestion that declination must be controlled by the speaker at least to some extent (e.g., Ladd, 1984). One of the arguments for declination being controlled

[3] Some studies, however, question the claim that the intonation phrase is a domain of declination and show that the domain of declination is sometimes larger than the intonation phrase; in other words, not all intonation phrases exhibit reset in fundamental frequency (Schuetze-Coburn *et al.*, 1991; cf. also Ladd's discussion of a hierarchy of declination units (1986)).

by higher level commands of the speaker, and thus being rather a phonological phenomenon, comes from the demonstration of a look-ahead mechanism in adjusting fundamental frequency range to the length of an utterance (O'Shaughnessy, 1976; Sorensen & Cooper, 1980; Paccia & Cooper, 1981). Another argument in favour of this view was provided by Grønnum (née Thorsen, 1980), who showed that declination differs according to sentence type, and by Vaissière (1983), who stated that declination is often suppressed in interrogative sentences.

With respect to interpreting a progressive decline in articulation rate within the intonation phrase, we have little insight. At this stage, one can only speculate about what causes the pattern to have its particular shape. To some extent, the same interpretations as suggested for the phrase-final lengthening may apply to a slowing down pattern as a whole, for I demonstrated that final lengthening is a part of this pattern. However, the origin of lengthening effect has not been conclusively established either (Bell-Berti, 1991; Bell-Berti *et al.*, 1995). I shall remark only on the main proposals.

One suggestion has been that lengthening is a part of the speaker's strategy to help the listener in decoding the message, particularly, by emphasizing phrase boundaries (Lindblom & Rapp, 1973; Lehiste, 1977; Lindblom *et al.*, 1981). It is easy to see that the same may apply to the whole pattern of articulation rate, i.e., also to the reset of rate at the beginning of the intonation phrase. This view would assume that articulation rate is planned at a higher level of speech production.

Another wide-spread belief is that final lengthening is produced in order to allow the speaker extra time for planning upcoming material in the next phrase (e.g., Cooper & Paccia-Cooper, 1980). One can speculate that a progressive decrease in articulation rate reflects ongoing planning of further material during the whole intonation phrase (the time when it was thought that planning of speech takes place only during pauses, and not while talking, is long gone; Brotherton, 1979;

Sabin et al., 1979). However, it is not possible to say what type of speech planning would occur, since this would depend on what parts of speech are planned.

Finally, it may just be that the rallentando pattern is a reflection of some sort of 'winding down', a relaxation response, as Bell-Berti *et al.* (1995) call it. They see it as a consequence of the speech system's natural rhythm. This, a more physiologically based explanation for decline in rate, would make it parallel to declination in fundamental frequency as seen by the proponents of the view of declination as an involuntary, automatic process. This would place rallentando on a very low level of speech production.

At present, the existing body of research is not sufficient to be able to decide between these views. The question at what level of speech production articulation rate patterning originates cannot yet be conclusively answered either. Only one language has been examined here, and the material is limited to declarative sentences. But the evidence so far suggests that the articulation rate pattern is generated during the stage of phonetic interpretation, i.e., at a low (phonetic) level of speech production, and that the shape of the pattern is not determined by higher levels of language structure (such as cognitive demands of a particular task, memory retrieval of the speech content, etc.). If some evidence is found for the existence of different patterns of articulation rate that would have a contrastive function (e.g., declarative vs. interrogative utterances), this would support a different interpretation, namely, that articulation rate variation is phonological and thus planned at a higher level of speech production. Other evidence in support of this view would be if the articulation rate at the beginning and end of an intonation phrase was independent of the number of words in the intonation phrase. This would mean that the slope of the decline in rate would be determined by the size of the intonation phrase as a whole, which would assume some sort of a look-ahead mechanism in a similar way as suggested for declination.

8.2 FACTORS AFFECTING ARTICULATION RATE OF PHONOLOGICAL WORDS

So far I have discussed the role of the position of a phonological word within the intonation phrase. However, in my thesis I showed that articulation rate of phonological words is partially determined also by their size in terms of the number of syllables and by the word class, i.e., whether they contain only function or content words.

With respect to the word's size, the tendency in both studies was that the more syllables the word contained, the faster was the rate with which it was produced. This tendency was significant in all three groups of intonation phrases, i.e., 2-word, 3-word, and 4-word phrases and applied to words in all the positions within the phrase. This is similar to what was found in other languages (e.g., Fletcher, 1991 for French; Williams & Hiller, 1994 for British English). An important and consistent feature of this relationship between words' articulation rate and their size is that it is not linear. The results of Study 2 suggest rather a power relationship; $AR = 4\ (sylls)^{0.4}$. So the rate starts levelling out for words with three or more syllables.

It was surprising to find that when the size of the word was expressed in terms of the number of segments (Study 1), the correlation between the articulation rate and size of the word disappeared. In this respect, it did not seem to be relevant whether long vowels counted as two segments or one.

In this thesis, two hypothetical explanations for the lack of correlation between word's articulation rate and its size in segments were suggested: (i) in Czech, polysyllabic words tend to contain syllables with a smaller number of consonants than monosyllabic or disyllabic words; in other words, they have a simpler syllabic structure, and/or (ii) some process of timing compensation operates within the syllable. In a further analysis I demonstrated that the former case applies to my data.

This does not necessarily mean that the number of segments in a word is completely irrelevant for the word's articulation rate. However, the effect would probably be seen only when words with the same syllabic complexity were compared. More analysis would be necessary to assess this effect. So far, it seems that in studies on articulation rate in Czech, the syllable is a more useful unit for measuring the phonological size of words than segments.

The issue of the relationship between the size of a word and its articulation rate is closely related to the issue of speech rhythm, in particular, stress-timing and syllable-timing. Traditionally, stress-timing (or isochrony) is supposed to be a tendency for stressed syllables to occur at equal intervals regardless of the number of intervening unstressed syllables. Syllable-timing is then a tendency for syllables to have equal durations, so inter-stress intervals tend to be longer in proportion to the number of syllables they contain (Abercrombie, 1967; Williams & Hiller, 1994). However, it is now well-known that most researchers seeking support for these two types of timing in the acoustic time domain failed to demonstrate persuasive evidence in their favour (e.g., Lehiste, 1973, 1975; Nakatani *et al.*, 1981; Roach, 1982; Wenk & Wioland, 1982; Dauer, 1983; Fant *et al.*, 1989; Fant *et al.*, 1991b; Fletcher, 1991). However, some researchers believe that a form of weak isochrony is present (e.g., Uldall, 1971) and that in connected speech a great number of timing factors operate, and they may easily override tendencies towards isochrony (Kohler, 1991).

Most studies on stress- vs. syllable-timing involved measuring the duration of interstress intervals; some also included mean syllable durations (e.g., Roach, 1982; Williams & Hiller, 1994). In this thesis, the analysis was based on articulation rate expressed in the number of syllables per second (which is just the reciprocal of mean syllable duration); duration of phonological words was not analysed directly. The finding for Czech that articulation rate increases with the number of syllables in the word would suggest that Czech has a tendency towards isochrony (in a purely syllable-timed language, and Czech used to be considered

as such, one would expect a constant articulation rate). However, the fact that the increase in rate gradually flattens in words with more than three syllables suggests that this tendency is rather weak.[4]

But there is a danger in drawing any strong conclusions about isochrony just on the basis of the articulation rate of phonological words. As Fletcher (1991) and Williams & Hiller (1994) warned, the lack of linear relationship between the number of syllables in the phonological word and articulation rate cannot be taken as evidence for isochrony; when they analysed the actual duration of the syllable or of the phonological word, they did find a linear relationship. A quick informal analysis of the data used in the present work confirmed that this applies to Czech as well (the two studies mentioned above used French and English respectively). As a possible explanation for this discrepancy, it was suggested that stressed syllables, which are lengthened, are included in the calculation and that, statistically speaking, the effect of this inclusion becomes less and less pronounced as words become larger (Fletcher, 1991). As an additional artefact, Williams & Hiller (1994) suggested that foot-initial syllables (i.e., stressed syllables) tend to have a larger number of consonants and a lower proportion of especially shorter vowels, e.g., schwa. The first suggestion is unlikely to apply in Czech, since stress does not lengthen syllable duration. The argument concerning the number of consonants in stressed syllables might be at least partly responsible for finding the same discrepancy in Czech; recall from the earlier discussion that polysyllabic words in Czech were found in this thesis to have a less complex syllable structure. A more detailed study would be necessary to make stronger conclusions about isochrony in Czech.

Returning now to the issue of determinants of articulation rate variation, another significant factor was word class. Out of the three groups of words, those

[4] Of course, purely syllable-timed and purely stress-timed languages can only be an abstraction; in reality, they are inconceivable, given the time constraints on articulatory movements. No language is likely to have five-syllable word pronounced five times faster than a monosyllable; but also no language is likely to have five-syllable word five times longer than a monosyllable!

containing only content words, those containing only function words and those which were a mixture of both, the most distinctly different with respect to articulation rate were the content and the function word categories. Unsurprisingly, the content word category was significantly slower than the function word category, but it was also slower than the mixed category. The former confirms what has been known about other languages (e.g., Kaiki *et al.*, 1990; O'Shaughnessy, 1995). The results were inconsistent as to whether the function word category and the mixed category differ in articulation rate. There were some indications that they do not, implying that the tendency for function words to be pronounced fast takes over. This would suggest that it might not be necessary to distinguish three levels within the factor of word class, and that it might be sufficient to have a binary classification, where the originally mixed level and the level of function words would be treated as a single level. However, when this proposition was put to test in Chapter 7, it did not improve the model of articulation rate, rather the contrary. Clearly, more research would be needed to settle this issue.

So far I have discussed linguistic factors affecting articulation rate variation in the data. However, the two extralinguistic factors, speaker and speech task also played some role, though not as strong as might have been expected, given the results reported in the literature. In the first study, based on a relatively small database, I found that only two speakers differed significantly in their articulation rate and only in intonation phrases containing four words. In the second study, examining a much larger dataset, the factor of speaker was significant overall. Again, however, not all subjects differed to the same extent; statistically speaking, only two out of four speakers were reliably different. These results cannot provide very strong support to what seems to be a general belief, that people differ in how fast they speak. Of course, there are speakers whom we notice as particularly slow or fast, but these are likely to be at the edge of the spectrum (Sedláková, 1989); for most speakers, their articulation rates do not differ to a significant extent, at least not at the level of phonological words.

More surprising perhaps was the finding that the three speech tasks, reading a story, paraphrasing the story and spontaneous speech, were not statistically different in articulation rate, with the exception of paraphrasing and spontaneous speech in three-word phrases, and reading and spontaneous speech in intonation phrases containing four words, where reading was significantly slower than spontaneous speech. This is completely contrary to what has been reported for other languages (e.g., for Finnish in Lehtonen, 1979; for German in Butcher, 1981; for Swedish in Strangert, 1993). Although this is an interesting finding, it cannot be generalized since it applied only to one group of phrases (which incidentally had the smallest number of observations). Overall, it can be concluded that there was no strong evidence in the data for task differences with respect to articulation rate.

Comparisons of the relative strength of the factors position of the word within the intonation phrase, its size in syllables, word class, speaker and task revealed that the first two had the strongest influence on articulation rate variation, with the remaining factors being far behind in their importance.

Finally, a quantitative description in the form of a linear equation was offered to predict articulation rate from the three linguistic factors. Six linear models in total were fitted to the data in Study 2. The best fitting model achieved 50% success in accounting for the rate variance, which is about the same as the overall success of the general linear model, which used all five factors, i.e., also speaker and task. The success in modelling is due to a logarithmic transformation of both articulation rate and the parameter of word size in syllables. Obviously, more factors are needed to explain a larger proportion of variance in articulation rate. Some of these are likely to be linguistic in nature. Possible candidates might be syntactic category of the word, whether it is focussed or not, the word's role in the information structure of the phrase, etc. It would also be worthwhile to investigate the relationship between the patterning of articulation rate and pause duration and location, for it has been suggested that articulation rate and pausing are related

I am not suggesting that research on articulation rate variation should be always based on the phonological word. The minimum unit for measurements could be larger; it depends on the purpose of the research (although I suggest that the two criteria stated above should be always kept in mind). As I mentioned in Chapter 1, most studies have used the interpause stretch, probably assuming that the pause is the place where articulation rate can be changed (e.g., Miller *et al.,* 1984; Eefting, 1990). The research presented in this thesis, especially if it is replicated for other languages, would suggest that if phrasal-level is considered, the unit to be used is the intonation phrase. It can be justified by (i) the fact that it specifies an underlying structure for the domain of articulation rate, and (ii) it is also sufficiently small to demonstrate variation in rate.

This brings me to the subject of the intonation phrase itself. Although listeners seem to intuitively identify intonation phrase boundaries with considerable agreement (e.g., de Pijper & Sanderman, 1995), for linguistic theory their exact perceptual correlates remain somewhat elusive (e.g., Schuetze-Coburn *et al.*, 1991; Selkirk, 1995; Ladd, 1996).

A number of prosodic cues have been identified which are used to both mark and identify intonation phrases. In general, intonation phrases are defined both in terms of their internal intonational structure and in terms of the location of phonetic boundaries (e.g., Ladd, 1986), or, as Cruttenden (1986) calls them, 'internal' and 'external' criteria. Internal features involve exclusively pitch characteristics, among them mainly 'coherent intonation contour' (Chafe 1987; Schuetze-Coburn *et al.*, 1991) and declination (Ladd, 1986). External features, or boundary markers, include both tonal characteristics (e.g., Stenström, 1990; Bruce *et al.*, 1991), such as phrase-initial pitch reset (Schuetze-Coburn *et al.*, 1991), pitch movement in the nucleus, if present, and in unstressed syllables following it (e.g., Cruttenden, 1986), and temporal characteristics, namely, the presence of a pause, syllable-final lengthening and acceleration of tempo on initial unstressed syllables (e.g., Cruttenden, 1986). Boundary markers may also involve some

aspects of voice quality, such as laryngealization (Chafe *et al.*, 1991; Knowles 1991).

It has been claimed that speakers use pauses, melodic markers and declination reset systematically to indicate prosodic boundaries (e.g., de Pijper & Sanderman, 1994; Sanderman & Collier, 1994; Sanderman 1995). Pitch movement and pauses seem to be used in similar ways across languages (Vincent, *et al.*, 1995), while final lengthening might be dependent on both language (Thorsen, 1980; Farnetani, 1989; Berkovits, 1984; Hallé *et al.*, 1991) and mode of discourse (Klatt 1975; Klatt & Cooper 1975).

The standard theory assumes that internal and external features co-occur. This does often happen, but not always (e.g., Ladd, 1986; Wichmann, 1995). The cases which do not involve both types of markers create a problem for linguistic theory, since the units derived by using external and internal criteria are thus not always the same. In practice, the assignment of intonation phrase boundaries thus often becomes rather a circular enterprise (Cruttenden, 1986; Ladd, 1986, 1996).

> "We establish some intonational groups [intonation phrases] in cases where all the external criteria conspire to make the assignment of a boundary relatively certain; we note the sorts of internal intonational structure occurring in such cases and this enables us to make decisions in those cases where the external criteria are less unambiguous." (Cruttenden, 1986:36)

A solution in terms of a hierarchy of two intonation units has been proposed (e.g., minor and major tone groups in Trim, 1959; intermediate phrase and intonation phrase in Beckman & Pierrehumbert, 1986; major phrase and tone group in Ladd, 1986).

Given the problems just described, it is not a great surprise that there is considerable disagreement as to the relative importance of the cues. For instance, Vincent *et al.* (1995) and Garrido *et al.* (1995) found that the strongest effect was

produced by phrase-final pitch drop and phrase-initial pitch reset, while neither pauses nor final lengthening seemed to matter. On the other hand, Schuetze-Coburn *et al.* (1991) note that, e.g., pitch reset is arguably more central than tempo modulation but that it is likely that both cues are required for an intonation phrase to be perceived. Yet another view is presented by Swerts & Geluykens (1994), who claim that the perception of a final boundary is carried out essentially on the basis of cues which precede the boundary. The results presented in this thesis suggest that not only phrase-final lengthening but also the phrase-internal structure of articulation rate, i.e., articulation rate reset, may also be used for identifying intonation phrases.

8.4 APPLICATIONS

From the speech technology point of view, research on articulation rate variation is highly relevant. As pointed out in the discussion of the role of prosody in speech technology, particularly, in speech synthesis and word recognition (Chapter 1), satisfactory modelling of various aspects of prosodic timing structure other than phrase-final lengthening is still lacking. Yet a number of studies suggested that introducing the factor of variation in articulation rate would contribute to improving both the naturalness of the signal and also the accuracy of automatic word recognition systems (e.g., Eefting, 1990; O'Shaughnessy, 1995). My work suggests that phrase-final lengthening is only part of a larger pattern. If this pattern turns out to be a linguistic universal, which is also perceptually relevant, both speech synthesis and word recognition would benefit from including it into their models. A step in this direction was made in this thesis by designing a mathematical model which could be used for predicting articulation rate on the basis of three linguistic factors, position of the word within the intonation phrase, its size in syllables and its word class.

Let me now turn to another aspect of speech technology, related to intonational analyses. Intonational analyses seem to assume that tonal and temporal cues for intonation phrases coincide. In these analyses, there is an apparent imbalance in attention paid to pitch and timing, although it is generally recognized that both are important with respect to intonation. For pitch, we have a clear physical measure in F_0 and well established categories for representation. For instance, the ToBI labelling system (Tones and Break Indices; Silverman *et al.*, 1992), which was developed to provide a standard transcription system of prosody for large corpora (originally for American English), and which is being used in a number of synthesis systems (Black & Taylor, 1994; Black & Hunt, 1996), uses categories of 'L' and 'H' tones. With respect to timing, there are two boundary phenomena, pauses and phrase-final lengthening, which are considered as representatives of timing. However, in intonation analyses, such as ToBI, timing is present only implicitly; there is no quantitative expression of time features.

However, there have been some attempts at semi-automatic quantification of articulation rate in continuous speech. An example of this is a study by Arnfield *et al.* (1995) who designed a system in which articulation rate is calculated from the number of syllables in the (lexical) word and its duration and then displayed within the ToBI labelling system. This way of quantifying articulation rate was aimed at detecting articulation rate variation in speech, and in this particular project, in emotional speech. If articulation rate patterning within the intonation phrase, as discovered in this thesis, turns out to apply also to English and other languages, it might represent additional, temporal information about the internal structure of the intonation phrase, expressed in terms of a timing (rate) contour, in a similar way as in Arnfield *et al.* (1995), but based on phonological words, rather than lexical words. Such a temporal representation could also help in identification of intonation phrases.

In conclusion, this thesis presents evidence that articulation rate in Czech does not vary in an arbitrary way but that its variation has a domain, the intonation phrase,

within which it follows a systematic pattern of progressive slowing down (rallentando). This pattern seems remarkably consistent both across speakers and across tasks. The methodology developed in this thesis can be applied to other languages; a unified approach would make it possible to compare findings for different languages. In fact, cross-linguistic studies are the most obvious extension of this work, and they would be essential for development of a comprehensive theory of articulation rate.

APPENDIX 1

Abbreviations:

SEGS1	word size in segments: long vowels counted as one segment
SEGS2	word size in segments: long vowels counted as two segments
SPEAKER	1 = MS
	2 = OL
	3 = JP

STUDY 1

GENERAL LINEAR MODELS

MODELS WITH MAIN FACTORS, AND 2-WAY AND 3-WAY INTERACTIONS (Saturated models)

Word size in segments (long vowel = 1 segment)

2-word IP's

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	182.423[b]	38	4.801	1.454	.121
Intercept	1557.602	1	1557.602	471.883	.000
POSITION	19.359	1	19.359	5.865	.020
SEGS1	15.822	9	1.758	.533	.842
SPEAKER	8.455	2	4.227	1.281	.289
POSITION * SEGS1	5.237	6	.873	.264	.950
POSITION * SPEAKER	12.053	2	6.026	1.826	.174
SEGS1 * SPEAKER	22.227	11	2.021	.612	.808
POSITION * SEGS1 * SPEAKER	8.368	6	1.395	.423	.860
Error	135.334	41	3.301		
Total	2950.971	80			
Corrected Total	317.757	79			

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Noncent. Parameter	Observed Power[a]
Corrected Model	55.266	.904
Intercept	471.883	1.000
POSITION	5.865	.657
SEGS1	4.793	.223
SPEAKER	2.561	.262
POSITION * SEGS1	1.587	.112
POSITION * SPEAKER	3.651	.359
SEGS1 * SPEAKER	6.734	.278
POSITION * SEGS1 * SPEAKER	2.535	.157
Error		
Total		
Corrected Total		

a. Computed using alpha = .05

b. R Squared = .574 (Adjusted R Squared = .179)

3-word IP's

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	287.139[b]	56	5.127	1.796	.011
Intercept	1805.689	1	1805.689	632.433	.000
POSITION	112.854	2	56.427	19.763	.000
SEGS1	30.002	12	2.500	.876	.575
SPEAKER	5.226	2	2.613	.915	.405
POSITION * SEGS1	37.951	13	2.919	1.022	.441
POSITION * SPEAKER	6.874	4	1.719	.602	.663
SEGS1 * SPEAKER	13.374	11	1.216	.426	.939
POSITION * SEGS1 * SPEAKER	27.133	10	2.713	.950	.494
Error	188.440	66	2.855		
Total	4661.777	123			
Corrected Total	475.578	122			

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Noncent. Parameter	Observed Power[a]
Corrected Model	100.569	.996
Intercept	632.433	1.000
POSITION	39.526	1.000
SEGS1	10.508	.456
SPEAKER	1.830	.202
POSITION * SEGS1	13.292	.553
POSITION * SPEAKER	2.408	.189
SEGS1 * SPEAKER	4.684	.209
POSITION * SEGS1 * SPEAKER	9.503	.453
Error		
Total		
Corrected Total		

a. Computed using alpha = .05

b. R Squared = .604 (Adjusted R Squared = .268)

4-word IP's

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	364.307[b]	63	5.783	2.443	.002
Intercept	2375.902	1	2375.902	1003.653	.000
POSITION	54.932	3	18.311	7.735	.000
SEGS1	15.879	8	1.985	.838	.575
SPEAKER	28.192	2	14.096	5.954	.006
POSITION * SEGS1	72.484	17	4.264	1.801	.068
POSITION * SPEAKER	21.724	6	3.621	1.529	.197
SEGS1 * SPEAKER	26.688	11	2.426	1.025	.446
POSITION * SEGS1 * SPEAKER	44.154	14	3.154	1.332	.237
Error	85.221	36	2.367		
Total	4427.580	100			
Corrected Total	449.528	99			

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Noncent. Parameter	Observed Power[a]
Corrected Model	153.894	.998
Intercept	1003.653	1.000
POSITION	23.205	.980
SEGS1	6.708	.326
SPEAKER	11.909	.851
POSITION * SEGS1	30.619	.851
POSITION * SPEAKER	9.177	.516
SEGS1 * SPEAKER	11.274	.460
POSITION * SEGS1 * SPEAKER	18.652	.650
Error		
Total		
Corrected Total		

a. Computed using alpha = .05

b. R Squared = .810 (Adjusted R Squared = .479)

Word size in segments (long vowel = 2 segments)

2-word IP's

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	205.164[b]	44	4.663	1.449	.130
Intercept	1454.757	1	1454.757	452.218	.000
POSITION	30.120	1	30.120	9.363	.004
SEGS2	24.989	10	2.499	.777	.650
SPEAKER	10.060	2	5.030	1.564	.224
POSITION * SEGS2	7.585	7	1.084	.337	.932
POSITION * SPEAKER	13.623	2	6.812	2.117	.135
SEGS2 * SPEAKER	36.929	14	2.638	.820	.644
POSITION * SEGS2 * SPEAKER	6.534	7	.933	.290	.953
Error	112.593	35	3.217		
Total	2950.971	80			
Corrected Total	317.757	79			

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Noncent. Parameter	Observed Power[a]
Corrected Model	63.776	.893
Intercept	452.218	1.000
POSITION	9.363	.845
SEGS2	7.768	.330
SPEAKER	3.127	.309
POSITION * SEGS2	2.358	.136
POSITION * SPEAKER	4.235	.405
SEGS2 * SPEAKER	11.479	.401
POSITION * SEGS2 * SPEAKER	2.031	.122
Error		
Total		
Corrected Total		

a. Computed using alpha = .05

b. R Squared = .646 (Adjusted R Squared = .200)

3-word IP's

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	342.178[b]	62	5.519	2.482	.000
Intercept	1969.761	1	1969.761	885.949	.000
POSITION	64.625	2	32.312	14.533	.000
SEGS2	35.572	12	2.964	1.333	.224
SPEAKER	8.610	2	4.305	1.936	.153
POSITION * SEGS2	57.372	14	4.098	1.843	.053
POSITION * SPEAKER	11.532	4	2.883	1.297	.282
SEGS2 * SPEAKER	40.193	14	2.871	1.291	.240
POSITION * SEGS2 * SPEAKER	45.290	12	3.774	1.698	.090
Error	133.400	60	2.223		
Total	4661.777	123			
Corrected Total	475.578	122			

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Noncent. Parameter	Observed Power[a]
Corrected Model	153.903	1.000
Intercept	885.949	1.000
POSITION	29.067	.998
SEGS2	15.999	.669
SPEAKER	3.873	.386
POSITION * SEGS2	25.804	.872
POSITION * SPEAKER	5.187	.381
SEGS2 * SPEAKER	18.078	.694
POSITION * SEGS2 * SPEAKER	20.370	.798
Error		
Total		
Corrected Total		

a. Computed using alpha = .05

b. R Squared = .719 (Adjusted R Squared = .430)

4-word IP's

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	349.135[b]	60	5.819	2.261	.004
Intercept	2196.683	1	2196.683	853.356	.000
POSITION	102.013	3	34.004	13.210	.000
SEGS2	4.746	9	.527	.205	.992
SPEAKER	29.448	2	14.724	5.720	.007
POSITION * SEGS2	66.716	18	3.706	1.440	.168
POSITION * SPEAKER	40.335	6	6.722	2.611	.032
SEGS2 * SPEAKER	36.311	10	3.631	1.411	.212
POSITION * SEGS2 * SPEAKER	44.410	11	4.037	1.568	.147
Error	100.393	39	2.574		
Total	4427.580	100			
Corrected Total	449.528	99			

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Noncent. Parameter	Observed Power[a]
Corrected Model	135.630	.997
Intercept	853.356	1.000
POSITION	39.630	1.000
SEGS2	1.844	.104
SPEAKER	11.440	.838
POSITION * SEGS2	25.917	.763
POSITION * SPEAKER	15.669	.792
SEGS2 * SPEAKER	14.106	.609
POSITION * SEGS2 * SPEAKER	17.252	.691
Error		
Total		
Corrected Total		

a. Computed using alpha = .05

b. R Squared = .777 (Adjusted R Squared = .433)

Word size in syllables

2-word IP's

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	157.317[b]	22	7.151	2.540	.003
Intercept	1017.095	1	1017.095	361.345	.000
POSITION	32.519	1	32.519	11.553	.001
SYLLS	36.986	5	7.397	2.628	.033
SPEAKER	5.765	2	2.883	1.024	.366
POSITION * SYLLS	1.008	2	.504	.179	.837
POSITION * SPEAKER	3.301	2	1.650	.586	.560
SYLLS * SPEAKER	8.342	4	2.085	.741	.568
POSITION * SYLLS * SPEAKER	2.275	4	.569	.202	.936
Error	160.441	57	2.815		
Total	2950.971	80			
Corrected Total	317.757	79			

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Noncent. Parameter	Observed Power[a]
Corrected Model	55.890	.992
Intercept	361.345	1.000
POSITION	11.553	.916
SYLLS	13.140	.766
SPEAKER	2.048	.220
POSITION * SYLLS	.358	.076
POSITION * SPEAKER	1.173	.143
SYLLS * SPEAKER	2.964	.224
POSITION * SYLLS * SPEAKER	.808	.090
Error		
Total		
Corrected Total		

a. Computed using alpha = .05

b. R Squared = .495 (Adjusted R Squared = .300)

3-word IP's

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	269.281[b]	33	8.160	3.520	.000
Intercept	1491.308	1	1491.308	643.375	.000
POSITION	47.091	2	23.546	10.158	.000
SYLLS	62.547	5	12.509	5.397	.000
SPEAKER	4.737	2	2.368	1.022	.364
POSITION * SYLLS	10.198	5	2.040	.880	.498
POSITION * SPEAKER	1.764	4	.441	.190	.943
SYLLS * SPEAKER	2.649	7	.378	.163	.992
POSITION * SYLLS * SPEAKER	19.041	7	2.720	1.173	.326
Error	206.297	89	2.318		
Total	4661.777	123			
Corrected Total	475.578	122			

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Noncent. Parameter	Observed Power[a]
Corrected Model	116.172	1.000
Intercept	643.375	1.000
POSITION	20.316	.984
SYLLS	26.984	.986
SPEAKER	2.043	.223
POSITION * SYLLS	4.399	.302
POSITION * SPEAKER	.761	.089
SYLLS * SPEAKER	1.143	.092
POSITION * SYLLS * SPEAKER	8.214	.478
Error		
Total		
Corrected Total		

a. Computed using alpha = .05

b. R Squared = .566 (Adjusted R Squared = .405)

4-word IP's

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	297.290[b]	37	8.035	3.272	.000
Intercept	1831.982	1	1831.982	746.090	.000
POSITION	57.419	3	19.140	7.795	.000
SYLLS	28.045	4	7.011	2.855	.031
SPEAKER	22.924	2	11.462	4.668	.013
POSITION * SYLLS	10.096	9	1.122	.457	.898
POSITION * SPEAKER	14.953	6	2.492	1.015	.424
SYLLS * SPEAKER	4.672	4	1.168	.476	.753
POSITION * SYLLS * SPEAKER	43.665	8	5.458	2.223	.037
Error	152.238	62	2.455		
Total	4427.580	100			
Corrected Total	449.528	99			

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Noncent. Parameter	Observed Power[a]
Corrected Model	121.074	1.000
Intercept	746.090	1.000
POSITION	23.385	.985
SYLLS	11.421	.744
SPEAKER	9.336	.765
POSITION * SYLLS	4.112	.205
POSITION * SPEAKER	6.090	.371
SYLLS * SPEAKER	1.903	.155
POSITION * SYLLS * SPEAKER	17.783	.822
Error		
Total		
Corrected Total		

a. Computed using alpha = .05

b. R Squared = .661 (Adjusted R Squared = .459)

APPENDIX 2

STUDY 1

GENERAL LINEAR MODELS

MODELS WITH THE MAIN FACTORS

Word size in segments (long vowel = 1 segment)

2-word IP's

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Corrected Model	113.836[b]	12	9.486	3.117	.001	37.402	.985
Intercept	1129.549	1	1129.549	371.123	.000	371.123	1.000
POSITION	77.151	1	77.151	25.349	.000	25.349	.999
SEGS1	17.157	9	1.906	.626	.771	5.637	.281
SPEAKER	17.651	2	8.826	2.900	.062	5.799	.549
Error	203.921	67	3.044				
Total	2950.971	80					
Corrected Total	317.757	79					

a. Computed using alpha = .05

b. R Squared = .358 (Adjusted R Squared = .243)

3-word IP's

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Corrected Model	194.405[b]	16	12.150	4.581	.000	73.289	1.000
Intercept	992.118	1	992.118	374.020	.000	374.020	1.000
POSITION	138.221	2	69.110	26.054	.000	52.108	1.000
SEGS1	33.432	12	2.786	1.050	.410	12.603	.576
SPEAKER	8.735	2	4.367	1.646	.198	3.293	.341
Error	281.173	106	2.653				
Total	4661.777	123					
Corrected Total	475.578	122					

a. Computed using alpha = .05

b. R Squared = .409 (Adjusted R Squared = .320)

Post Hoc Tests

POSITION

Multiple Comparisons

Dependent Variable: AR
Tukey HSD

(I) POSITION	(J) POSITION	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	-1.0390E-02	.360	1.000	-.8655	.8447
	3.00	2.3225*	.360	.000	1.4674	3.1776
2.00	1.00	1.039E-02	.360	1.000	-.8447	.8655
	3.00	2.3329*	.360	.000	1.4778	3.1880
3.00	1.00	-2.3225*	.360	.000	-3.1776	-1.4674
	2.00	-2.3329*	.360	.000	-3.1880	-1.4778

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR

Tukey HSD[a,b]

POSITION	N	Subset 1	Subset 2
3.00	41	4.2821	
1.00	41		6.6046
2.00	41		6.6150
Sig.		1.000	1.000

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 2.653.

a. Uses Harmonic Mean Sample Size = 41.000.

b. Alpha = .05.

4-word IP's

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Corrected Model	194.775[b]	13	14.983	5.058	.000	65.753	1.000
Intercept	1745.580	1	1745.580	589.278	.000	589.278	1.000
POSITION	100.601	3	33.534	11.320	.000	33.961	.999
SEGS1	16.481	8	2.060	.695	.695	5.564	.302
SPEAKER	63.940	2	31.970	10.792	.000	21.585	.988
Error	254.752	86	2.962				
Total	4427.580	100					
Corrected Total	449.528	99					

a. Computed using alpha = .05

b. R Squared = .433 (Adjusted R Squared = .348)

Post Hoc Tests

POSITION

Multiple Comparisons

Dependent Variable: AR
Tukey HSD

(I) POSITION	(J) POSITION	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	.2945	.487	.930	-.9809	1.5699
	3.00	.8186	.487	.340	-.4568	2.0940
	4.00	2.7577*	.487	.000	1.4823	4.0331
2.00	1.00	-.2945	.487	.930	-1.5699	.9809
	3.00	.5241	.487	.705	-.7513	1.7995
	4.00	2.4632*	.487	.000	1.1877	3.7386
3.00	1.00	-.8186	.487	.340	-2.0940	.4568
	2.00	-.5241	.487	.705	-1.7995	.7513
	4.00	1.9391*	.487	.001	.6637	3.2145
4.00	1.00	-2.7577*	.487	.000	-4.0331	-1.4823
	2.00	-2.4632*	.487	.000	-3.7386	-1.1877
	3.00	-1.9391*	.487	.001	-3.2145	-.6637

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR

Tukey HSD[a,b]

POSITION	N	Subset 1	Subset 2
4.00	25	4.5172	
3.00	25		6.4563
2.00	25		6.9804
1.00	25		7.2749
Sig.		1.000	.340

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 2.962.

a. Uses Harmonic Mean Sample Size = 25.000.

b. Alpha = .05.

SPEAKER

Multiple Comparisons

Dependent Variable: AR
Tukey HSD

(I) SPEAKER	(J) SPEAKER	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	1.1026*	.460	.049	5.518E-03	2.1996
	3.00	-.8132	.416	.130	-1.8056	.1791
2.00	1.00	-1.1026*	.460	.049	-2.1996	-5.5179E-03
	3.00	-1.9158*	.416	.000	-2.9081	-.9235
3.00	1.00	.8132	.416	.130	-.1791	1.8056
	2.00	1.9158*	.416	.000	.9235	2.9081

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR

Tukey HSD[a,b,c]

SPEAKER	N	Subset 1	Subset 2
2.00	28	5.1555	
1.00	28		6.2581
3.00	44		7.0713
Sig.		1.000	.149

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 2.962.

a. Uses Harmonic Mean Sample Size = 31.862.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

c. Alpha = .05.

Word size in segments (long vowel = 2 segments)

2-word IP's

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Corrected Model	126.422[b]	13	9.725	3.355	.001	43.609	.993
Intercept	998.470	1	998.470	344.417	.000	344.417	1.000
POSITION	76.355	1	76.355	26.338	.000	26.338	.999
SEGS2	29.742	10	2.974	1.026	.432	10.259	.489
SPEAKER	17.710	2	8.855	3.054	.054	6.109	.572
Error	191.335	66	2.899				
Total	2950.971	80					
Corrected Total	317.757	79					

a. Computed using alpha = .05

b. R Squared = .398 (Adjusted R Squared = .279)

3-word IP's

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Corrected Model	207.372[b]	16	12.961	5.122	.000	81.957	1.000
Intercept	1177.400	1	1177.400	465.329	.000	465.329	1.000
POSITION	134.506	2	67.253	26.580	.000	53.159	1.000
SEGS2	46.399	12	3.867	1.528	.125	18.338	.780
SPEAKER	7.697	2	3.849	1.521	.223	3.042	.317
Error	268.207	106	2.530				
Total	4661.777	123					
Corrected Total	475.578	122					

a. Computed using alpha = .05

b. R Squared = .436 (Adjusted R Squared = .351)

Post Hoc Tests

POSITION

Multiple Comparisons

Dependent Variable: AR
Tukey HSD

(I) POSITION	(J) POSITION	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	-1.0390E-02	.351	1.000	-.8455	.8247
	3.00	2.3225*	.351	.000	1.4874	3.1576
2.00	1.00	1.039E-02	.351	1.000	-.8247	.8455
	3.00	2.3329*	.351	.000	1.4978	3.1680
3.00	1.00	-2.3225*	.351	.000	-3.1576	-1.4874
	2.00	-2.3329*	.351	.000	-3.1680	-1.4978

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR

Tukey HSD[a,b]

POSITION	N	Subset 1	Subset 2
3.00	41	4.2821	
1.00	41		6.6046
2.00	41		6.6150
Sig.		1.000	1.000

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 2.530.

a. Uses Harmonic Mean Sample Size = 41.000.

b. Alpha = .05.

4-word IP's

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Corrected Model	188.197[b]	14	13.443	4.372	.000	61.213	1.000
Intercept	1599.800	1	1599.800	520.348	.000	520.348	1.000
POSITION	101.702	3	33.901	11.026	.000	33.079	.999
SEGS2	9.902	9	1.100	.358	.952	3.221	.169
SPEAKER	49.107	2	24.554	7.986	.001	15.973	.950
Error	261.331	85	3.074				
Total	4427.580	100					
Corrected Total	449.528	99					

a. Computed using alpha = .05

b. R Squared = .419 (Adjusted R Squared = .323)

Post Hoc Tests

POSITION

Multiple Comparisons

Dependent Variable: AR
Tukey HSD

(I) POSITION	(J) POSITION	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	.2945	.496	.934	-1.0051	1.5942
	3.00	.8186	.496	.356	-.4811	2.1183
	4.00	2.7577*	.496	.000	1.4580	4.0573
2.00	1.00	-.2945	.496	.934	-1.5942	1.0051
	3.00	.5241	.496	.717	-.7756	1.8237
	4.00	2.4632*	.496	.000	1.1635	3.7628
3.00	1.00	-.8186	.496	.356	-2.1183	.4811
	2.00	-.5241	.496	.717	-1.8237	.7756
	4.00	1.9391*	.496	.001	.6394	3.2387
4.00	1.00	-2.7577*	.496	.000	-4.0573	-1.4580
	2.00	-2.4632*	.496	.000	-3.7628	-1.1635
	3.00	-1.9391*	.496	.001	-3.2387	-.6394

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR

Tukey HSD[a,b]

POSITION	N	Subset 1	Subset 2
4.00	25	4.5172	
3.00	25		6.4563
2.00	25		6.9804
1.00	25		7.2749
Sig.		1.000	.356

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 3.074.

a. Uses Harmonic Mean Sample Size = 25.000.

b. Alpha = .05.

SPEAKER

Multiple Comparisons

Dependent Variable: AR
Tukey HSD

(I) SPEAKER	(J) SPEAKER	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	1.1026	.469	.054	-1.5301E-02	2.2204
	3.00	-.8132	.424	.140	-1.8244	.1979
2.00	1.00	-1.1026	.469	.054	-2.2204	1.530E-02
	3.00	-1.9158*	.424	.000	-2.9270	-.9047
3.00	1.00	.8132	.424	.140	-.1979	1.8244
	2.00	1.9158*	.424	.000	.9047	2.9270

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR

Tukey HSD[a,b,c]

SPEAKER	N	Subset 1	Subset 2
2.00	28	5.1555	
1.00	28		6.2581
3.00	44		7.0713
Sig.		1.000	.159

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 3.074.

a. Uses Harmonic Mean Sample Size = 31.862.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

c. Alpha = .05.

Word size in syllables

2-word IP's

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Corrected Model	127.455[b]	7	18.208	7.183	.000	50.278	1.000
Intercept	1045.524	1	1045.524	412.436	.000	412.436	1.000
POSITION	62.690	1	62.690	24.730	.000	24.730	.998
SYLLS	37.746	4	9.437	3.723	.008	14.890	.864
SPEAKER	9.712	2	4.856	1.916	.155	3.831	.385
Error	179.985	71	2.535				
Total	2871.244	79					
Corrected Total	307.440	78					

a. Computed using alpha = .05

b. R Squared = .415 (Adjusted R Squared = .357)

Post Hoc Tests

SYLLS

Multiple Comparisons

Dependent Variable: AR
Tukey HSD

(I) SYLLS	(J) SYLLS	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	-1.2972	.959	.659	-3.9814	1.3870
	3.00	-2.4919	.971	.088	-5.2094	.2255
	4.00	-2.5385	1.020	.105	-5.3931	.3161
	5.00	-4.3790*	1.300	.010	-8.0179	-.7401
2.00	1.00	1.2972	.959	.659	-1.3870	3.9814
	3.00	-1.1947*	.415	.041	-2.3558	-3.3653E-02
	4.00	-1.2413	.519	.130	-2.6946	.2120
	5.00	-3.0818*	.959	.016	-5.7660	-.3976
3.00	1.00	2.4919	.971	.088	-.2255	5.2094
	2.00	1.1947*	.415	.041	3.365E-02	2.3558
	4.00	-4.6577E-02	.541	1.000	-1.5604	1.4673
	5.00	-1.8871	.971	.304	-4.6045	.8304
4.00	1.00	2.5385	1.020	.105	-.3161	5.3931
	2.00	1.2413	.519	.130	-.2120	2.6946
	3.00	4.658E-02	.541	1.000	-1.4673	1.5604
	5.00	-1.8405	1.020	.379	-4.6951	1.0141
5.00	1.00	4.3790*	1.300	.010	.7401	8.0179
	2.00	3.0818*	.959	.016	.3976	5.7660
	3.00	1.8871	.971	.304	-.8304	4.6045
	4.00	1.8405	1.020	.379	-1.0141	4.6951

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR

Tukey HSD[a,b,c]

SYLLS	N	Subset 1	Subset 2
1.00	3	3.7343	
2.00	34	5.0315	
3.00	26	6.2263	6.2263
4.00	13	6.2728	6.2728
5.00	3		8.1133
Sig.		.050	.240

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 2.535.

a. Uses Harmonic Mean Sample Size = 6.162.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

c. Alpha = .05.

3-word IP's

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Corrected Model	229.331[b]	8	28.666	13.158	.000	105.265	1.000
Intercept	1776.721	1	1776.721	815.527	.000	815.527	1.000
POSITION	119.133	2	59.566	27.341	.000	54.683	1.000
SYLLS	68.319	4	17.080	7.840	.000	31.359	.997
SPEAKER	15.882	2	7.941	3.645	.029	7.290	.662
Error	246.184	113	2.179				
Total	4624.750	122					
Corrected Total	475.515	121					

a. Computed using alpha = .05

b. R Squared = .482 (Adjusted R Squared = .446)

Post Hoc Tests

POSITION

Multiple Comparisons

Dependent Variable: AR
Tukey HSD

(I) POSITION	(J) POSITION	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	2.599E-03	.328	1.000	-.7765	.7817
	3.00	2.3355*	.328	.000	1.5564	3.1146
2.00	1.00	-2.5994E-03	.328	1.000	-.7817	.7765
	3.00	2.3329*	.326	.000	1.5587	3.1071
3.00	1.00	-2.3355*	.328	.000	-3.1146	-1.5564
	2.00	-2.3329*	.326	.000	-3.1071	-1.5587

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR

Tukey HSD[a,b,c]

POSITION	N	Subset 1	Subset 2
3.00	41	4.2821	
2.00	41		6.6150
1.00	40		6.6176
Sig.		1.000	1.000

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 2.179.

a. Uses Harmonic Mean Sample Size = 40.661.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

c. Alpha = .05.

SYLLS

Multiple Comparisons

Dependent Variable: AR
Tukey HSD

(I) SYLLS	(J) SYLLS	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	-2.4556*	.636	.002	-4.2199	-.6913
	3.00	-2.8218*	.642	.000	-4.6026	-1.0410
	4.00	-3.3231*	.713	.000	-5.2998	-1.3465
	5.00	-5.5275*	.894	.000	-8.0053	-3.0496
2.00	1.00	2.4556*	.636	.002	.6913	4.2199
	3.00	-.3662	.302	.745	-1.2044	.4720
	4.00	-.8675	.433	.270	-2.0668	.3318
	5.00	-3.0718*	.691	.000	-4.9878	-1.1559
3.00	1.00	2.8218*	.642	.000	1.0410	4.6026
	2.00	.3662	.302	.745	-.4720	1.2044
	4.00	-.5013	.441	.787	-1.7248	.7222
	5.00	-2.7056*	.697	.002	-4.6368	-.7745
4.00	1.00	3.3231*	.713	.000	1.3465	5.2998
	2.00	.8675	.433	.270	-.3318	2.0668
	3.00	.5013	.441	.787	-.7222	1.7248
	5.00	-2.2043*	.762	.036	-4.3174	-9.1223E-02
5.00	1.00	5.5275*	.894	.000	3.0496	8.0053
	2.00	3.0718*	.691	.000	1.1559	4.9878
	3.00	2.7056*	.697	.002	.7745	4.6368
	4.00	2.2043*	.762	.036	9.122E-02	4.3174

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR

Tukey HSD[a,b,c]

SYLLS	N	Subset 1	Subset 2	Subset 3
1.00	6	3.1323		
2.00	52		5.5880	
3.00	44		5.9542	
4.00	15		6.4555	
5.00	5			8.6598
Sig.		1.000	.662	1.000

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 2.179.

a. Uses Harmonic Mean Sample Size = 10.520.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

c. Alpha = .05.

SPEAKER

Multiple Comparisons

Dependent Variable: AR
Tukey HSD

(I) SPEAKER	(J) SPEAKER	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1.00	2.00	5.890E-02	.338	.983	-.7434	.8612
	3.00	-.6433	.333	.134	-1.4334	.1467
2.00	1.00	-5.8895E-02	.338	.983	-.8612	.7434
	3.00	-.7022	.317	.073	-1.4543	4.988E-02
3.00	1.00	.6433	.333	.134	-.1467	1.4334
	2.00	.7022	.317	.073	-4.9882E-02	1.4543

Based on observed means. The error term is Error.

Homogeneous Subsets

AR

Tukey HSD[a,b,c]

SPEAKER	N	Subset 1
2.00	42	5.5559
1.00	35	5.6148
3.00	45	6.2581
Sig.		.088

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 2.179.

a. Uses Harmonic Mean Sample Size = 40.213.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

c. Alpha = .05.

4-word IP's

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Corrected Model	212.497[b]	9	23.611	8.965	.000	80.685	1.000
Intercept	1492.885	1	1492.885	566.846	.000	566.846	1.000
POSITION	94.116	3	31.372	11.912	.000	35.736	1.000
SYLLS	34.203	4	8.551	3.247	.015	12.987	.814
SPEAKER	44.428	2	22.214	8.435	.000	16.869	.960
Error	237.030	90	2.634				
Total	4427.580	100					
Corrected Total	449.528	99					

a. Computed using alpha = .05

b. R Squared = .473 (Adjusted R Squared = .420)

Post Hoc Tests

POSITION

Multiple Comparisons

Dependent Variable: AR
Tukey HSD

(I) POSITION	(J) POSITION	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	.2945	.459	.918	-.9070	1.4961
	3.00	.8186	.459	.288	-.3830	2.0202
	4.00	2.7577*	.459	.000	1.5561	3.9592
2.00	1.00	-.2945	.459	.918	-1.4961	.9070
	3.00	.5241	.459	.665	-.6775	1.7256
	4.00	2.4632*	.459	.000	1.2616	3.6647
3.00	1.00	-.8186	.459	.288	-2.0202	.3830
	2.00	-.5241	.459	.665	-1.7256	.6775
	4.00	1.9391*	.459	.000	.7375	3.1406
4.00	1.00	-2.7577*	.459	.000	-3.9592	-1.5561
	2.00	-2.4632*	.459	.000	-3.6647	-1.2616
	3.00	-1.9391*	.459	.000	-3.1406	-.7375

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR

Tukey HSD[a,b]

POSITION	N	Subset 1	Subset 2
4.00	25	4.5172	
3.00	25		6.4563
2.00	25		6.9804
1.00	25		7.2749
Sig.		1.000	.288

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 2.634.

a. Uses Harmonic Mean Sample Size = 25.000.

b. Alpha = .05.

SYLLS

Multiple Comparisons

Dependent Variable: AR
Tukey HSD

(I) SYLLS	(J) SYLLS	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	-3.7640*	.849	.000	-6.1281	-1.4000
	3.00	-3.6753*	.853	.000	-6.0501	-1.3005
	4.00	-4.9887*	.937	.000	-7.5971	-2.3804
	5.00	-5.0130*	1.148	.000	-8.2076	-1.8184
2.00	1.00	3.7640*	.849	.000	1.4000	6.1281
	3.00	8.871E-02	.363	.999	-.9228	1.1002
	4.00	-1.2247	.531	.153	-2.7035	.2541
	5.00	-1.2490	.849	.584	-3.6130	1.1151
3.00	1.00	3.6753*	.853	.000	1.3005	6.0501
	2.00	-8.8708E-02	.363	.999	-1.1002	.9228
	4.00	-1.3134	.537	.113	-2.8094	.1826
	5.00	-1.3377	.853	.522	-3.7125	1.0371
4.00	1.00	4.9887*	.937	.000	2.3804	7.5971
	2.00	1.2247	.531	.153	-.2541	2.7035
	3.00	1.3134	.537	.113	-.1826	2.8094
	5.00	-2.4250E-02	.937	1.000	-2.6326	2.5841
5.00	1.00	5.0130*	1.148	.000	1.8184	8.2076
	2.00	1.2490	.849	.584	-1.1151	3.6130
	3.00	1.3377	.853	.522	-1.0371	3.7125
	4.00	2.425E-02	.937	1.000	-2.5841	2.6326

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR

Tukey HSD[a,b,c]

SYLLS	N	Subset 1	Subset 2
1.00	4	2.5305	
3.00	38		6.2058
2.00	42		6.2945
4.00	12		7.5193
5.00	4		7.5435
Sig.		1.000	.478

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 2.634.

a. Uses Harmonic Mean Sample Size = 7.893.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

c. Alpha = .05.

SPEAKER

Multiple Comparisons

Dependent Variable: AR
Tukey HSD

(I) SPEAKER	(J) SPEAKER	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1.00	2.00	1.1026*	.434	.034	6.895E-02	2.1362
	3.00	-.8132	.392	.101	-1.7482	.1217
2.00	1.00	-1.1026*	.434	.034	-2.1362	-6.8949E-02
	3.00	-1.9158*	.392	.000	-2.8508	-.9809
3.00	1.00	.8132	.392	.101	-.1217	1.7482
	2.00	1.9158*	.392	.000	.9809	2.8508

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR

Tukey HSD[a,b,c]

SPEAKER	N	Subset 1	Subset 2
2.00	28	5.1555	
1.00	28		6.2581
3.00	44		7.0713
Sig.		1.000	.118

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 2.634.

a. Uses Harmonic Mean Sample Size = 31.862.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

c. Alpha = .05.

APPENDIX 3

STUDY 1

GENERAL LINEAR MODELS

PATTERN ABC

Speaker MS

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	49.594[b]	2	24.797	20.969	.000
Intercept	642.929	1	642.929	543.686	.000
POSITION	49.594	2	24.797	20.969	.000
Error	21.286	18	1.183		
Total	713.808	21			
Corrected Total	70.879	20			

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Noncent. Parameter	Observed Power[a]
Corrected Model	41.938	1.000
Intercept	543.686	1.000
POSITION	41.938	1.000
Error		
Total		
Corrected Total		

a. Computed using alpha = .05

b. R Squared = .700 (Adjusted R Squared = .666)

Post Hoc Tests

POSITION

Multiple Comparisons

Dependent Variable: AR
Tukey HSD

(I) POSITION	(J) POSITION	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	1.9973*	.581	.008	.5138	3.4808
	3.00	3.7619*	.581	.000	2.2784	5.2453
2.00	1.00	-1.9973*	.581	.008	-3.4808	-.5138
	3.00	1.7646*	.581	.019	.2811	3.2481
3.00	1.00	-3.7619*	.581	.000	-5.2453	-2.2784
	2.00	-1.7646*	.581	.019	-3.2481	-.2811

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR

Tukey HSD[a,b]

POSITION	N	Subset		
		1	2	3
3.00	7	3.6910		
2.00	7		5.4556	
1.00	7			7.4529
Sig.		1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 1.183.

a. Uses Harmonic Mean Sample Size = 7.000.

b. Alpha = .05.

Speaker OL

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	21.709[b]	2	10.855	7.666	.007
Intercept	484.618	1	484.618	342.259	.000
POSITION	21.709	2	10.855	7.666	.007
Error	16.991	12	1.416		
Total	523.318	15			
Corrected Total	38.701	14			

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Noncent. Parameter	Observed Power[a]
Corrected Model	15.332	.877
Intercept	342.259	1.000
POSITION	15.332	.877
Error		
Total		
Corrected Total		

a. Computed using alpha = .05

b. R Squared = .561 (Adjusted R Squared = .488)

Oxford University Phonetics Laboratory

Post Hoc Tests

POSITION

Multiple Comparisons

Dependent Variable: AR
Tukey HSD

(I) POSITION	(J) POSITION	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	1.4638	.753	.169	-.5440	3.4716
	3.00	2.9468*	.753	.005	.9390	4.9546
2.00	1.00	-1.4638	.753	.169	-3.4716	.5440
	3.00	1.4830	.753	.162	-.5248	3.4908
3.00	1.00	-2.9468*	.753	.005	-4.9546	-.9390
	2.00	-1.4830	.753	.162	-3.4908	.5248

Based on observed means. The error term is Error.
*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR

Tukey HSD[a,b]

POSITION	N	Subset 1	Subset 2
3.00	5	4.2074	
2.00	5	5.6904	5.6904
1.00	5		7.1542
Sig.		.162	.169

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 1.416.
a. Uses Harmonic Mean Sample Size = 5.000.
b. Alpha = .05.

Speaker JP

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	43.223[b]	2	21.612	13.858	.000
Intercept	1065.933	1	1065.933	683.528	.000
POSITION	43.223	2	21.612	13.858	.000
Error	32.749	21	1.559		
Total	1141.905	24			
Corrected Total	75.972	23			

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Noncent. Parameter	Observed Power[a]
Corrected Model	27.717	.995
Intercept	683.528	1.000
POSITION	27.717	.995
Error		
Total		
Corrected Total		

a. Computed using alpha = .05

b. R Squared = .569 (Adjusted R Squared = .528)

Post Hoc Tests

POSITION

Multiple Comparisons

Dependent Variable: AR
Tukey HSD

(I) POSITION	(J) POSITION	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	1.4416	.624	.077	-.1322	3.0155
	3.00	3.2792*	.624	.000	1.7054	4.8531
2.00	1.00	-1.4416	.624	.077	-3.0155	.1322
	3.00	1.8376*	.624	.020	.2638	3.4115
3.00	1.00	-3.2792*	.624	.000	-4.8531	-1.7054
	2.00	-1.8376*	.624	.020	-3.4115	-.2638

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR

Tukey HSD[a,b]

POSITION	N	Subset 1	Subset 2
3.00	8	4.9588	
2.00	8		6.7964
1.00	8		8.2380
Sig.		1.000	.077

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 1.559.

a. Uses Harmonic Mean Sample Size = 8.000.

b. Alpha = .05.

PATTERN BAC

Speaker MS

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	25.462[b]	2	12.731	12.649	.002
Intercept	386.944	1	386.944	384.454	.000
POSITION	25.462	2	12.731	12.649	.002
Error	9.058	9	1.006		
Total	421.465	12			
Corrected Total	34.521	11			

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Noncent. Parameter	Observed Power[a]
Corrected Model	25.298	.971
Intercept	384.454	1.000
POSITION	25.298	.971
Error		
Total		
Corrected Total		

a. Computed using alpha = .05

b. R Squared = .738 (Adjusted R Squared = .679)

Post Hoc Tests

POSITION

Multiple Comparisons

Dependent Variable: AR
Tukey HSD

(I) POSITION	(J) POSITION	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	-.9743	.709	.394	-2.9549	1.0064
	3.00	2.4855*	.709	.017	.5049	4.4661
2.00	1.00	.9743	.709	.394	-1.0064	2.9549
	3.00	3.4598*	.709	.002	1.4791	5.4404
3.00	1.00	-2.4855*	.709	.017	-4.4661	-.5049
	2.00	-3.4598*	.709	.002	-5.4404	-1.4791

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR

Tukey HSD[a,b]

POSITION	N	Subset 1	Subset 2
3.00	4	3.6968	
1.00	4		6.1823
2.00	4		7.1565
Sig.		1.000	.394

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 1.006.

a. Uses Harmonic Mean Sample Size = 4.000.

b. Alpha = .05.

Speaker OL

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	39.068[b]	2	19.534	5.822	.013
Intercept	540.372	1	540.372	161.055	.000
POSITION	39.068	2	19.534	5.822	.013
Error	50.328	15	3.355		
Total	629.768	18			
Corrected Total	89.396	17			

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Noncent. Parameter	Observed Power[a]
Corrected Model	11.644	.793
Intercept	161.055	1.000
POSITION	11.644	.793
Error		
Total		
Corrected Total		

a. Computed using alpha = .05

b. R Squared = .437 (Adjusted R Squared = .362)

Post Hoc Tests

POSITION

Multiple Comparisons

Dependent Variable: AR
Tukey HSD

(I) POSITION	(J) POSITION	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	-1.7232	1.058	.264	-4.4701	1.0238
	3.00	1.8843	1.058	.209	-.8626	4.6313
2.00	1.00	1.7232	1.058	.264	-1.0238	4.4701
	3.00	3.6075*	1.058	.010	.8606	6.3544
3.00	1.00	-1.8843	1.058	.209	-4.6313	.8626
	2.00	-3.6075*	1.058	.010	-6.3544	-.8606

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR

Tukey HSD[a,b]

POSITION	N	Subset 1	Subset 2
3.00	6	3.6485	
1.00	6	5.5328	5.5328
2.00	6		7.2560
Sig.		.209	.264

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 3.355.

a. Uses Harmonic Mean Sample Size = 6.000.

b. Alpha = .05.

Speaker JP

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	47.440[b]	2	23.720	5.644	.015
Intercept	655.811	1	655.811	156.043	.000
POSITION	47.440	2	23.720	5.644	.015
Error	63.041	15	4.203		
Total	766.293	18			
Corrected Total	110.481	17			

Tests of Between-Subjects Effects

Dependent Variable: AR

Source	Noncent. Parameter	Observed Power[a]
Corrected Model	11.288	.779
Intercept	156.043	1.000
POSITION	11.288	.779
Error		
Total		
Corrected Total		

a. Computed using alpha = .05

b. R Squared = .429 (Adjusted R Squared = .353)

Post Hoc Tests

POSITION

Multiple Comparisons

Dependent Variable: AR
Tukey HSD

(I) POSITION	(J) POSITION	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	-2.3017	1.184	.161	-5.3761	.7727
	3.00	1.6575	1.184	.366	-1.4169	4.7319
2.00	1.00	2.3017	1.184	.161	-.7727	5.3761
	3.00	3.9592*	1.184	.012	.8848	7.0336
3.00	1.00	-1.6575	1.184	.366	-4.7319	1.4169
	2.00	-3.9592*	1.184	.012	-7.0336	-.8848

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR

Tukey HSD[a,b]

POSITION	N	Subset 1	Subset 2
3.00	6	4.1638	
1.00	6	5.8213	5.8213
2.00	6		8.1230
Sig.		.366	.161

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 4.203.

a. Uses Harmonic Mean Sample Size = 6.000.

b. Alpha = .05.

APPENDIX 4

STUDY 1

ONE-WAY ANOVA'S
(significant results only)

Factor word size in syllables

Speaker MS (2-word phrases)

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
AR	Between Groups	43.006	4	10.752	8.102	.003
	Within Groups	14.598	11	1.327		
	Total	57.604	15			

Speaker OL (3-word phrases)

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
AR	Between Groups	40.765	4	10.191	3.630	.014
	Within Groups	103.872	37	2.807		
	Total	144.637	41			

Speaker OL (4-word phrases)

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
AR	Between Groups	44.478	3	14.826	7.722	.001
	Within Groups	46.078	24	1.920		
	Total	90.556	27			

Factor word size in segments

Speaker MS (2-word phrases)

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
AR	Between Groups	47.518	6	7.920	7.067	.005
	Within Groups	10.086	9	1.121		
	Total	57.604	15			

APPENDIX 5

Reading passage used in the analysis in Study 2

Když jsem byl od Huga asi tři metry, říkám mu, co tu dělá. Jenom něco zamumlal. A tak divně se zasmál a pak ustupoval ode mě a mával přitom rukama jako kachna, která se ubírá pryč mělčinou. Najednou mě popadla hrůza. Vzpomněl jsem si na horoskop, který před lety předpověděl, že se někdo z naší rodiny jednou zblázní, a napadlo mě, že se Hugo zbláznil. Chechtal se a šel tou vodou dál a mířil k hloubce. Skočil jsem do stanu, třepal s Jirkou a přitom křičel, že Hugo se zbláznil a že se utopí. Jirka vyskočil, spal nahý, a vlítli jsme na kraj řeky. Jirka byl od Huga asi tři metry, když se ke mně otočil a pronesl důrazně a káravě, že to není Hugo. V té chvíli vám proběhne hlavou tisíc věcí. První mě napadlo, že je to zloděj anebo vrah. Jirka byl u něho blíž než já a to bylo dobře, protože Jirka se nikdy ničeho a nikoho nebál, stejně jako můj tatínek. Viděl jsem ho, jak srazil daleko větší chlapy k zemi, když mu ublížili anebo ublížili tomu, za co on se bil. Proto bylo dobré, že byl první. Já jsem se bál. Postupoval k tomu chlapovi a ten chlap, který byl mohutný a podobal se skutečně Hugovi, čekal. Všiml jsem si, že má ruku pod vodou. Začal jsem se bát ještě víc a řekl jsem tiše, že může mít v ruce sekeru. Jirku to zastavilo. Stáli tam v řece proti sobě dlouho. Jirka neustupoval, ani jeden z nich nechtěl být zřejmě zbabělý. A pak ten chlap udělal dva tři kroky zpátky, zas se tak divně zasmál a odplaval na druhou stranu řeky. Na protějším břehu zmizel zřejmě v křoví. Šli jsme vzbudit rybáře. Tam, kde stály japonské pruty, bylo prázdné místo.

English translation (by James Naughton)

When I was about three meters away from him, I asked him what he was doing there. He mumbled something like: oh, you know. And then he gave a sort of funny laugh and backed away from me, flapping his arms like a duck moving off through the shallows. Suddenly, I was struck with horror. I remembered the horoscope which years ago had predicted that one day someone in our family would go mad, and it occurred to me that Hugo had gone mad. Chuckling, he waded on through the water, making for the depths. I leapt into the tent, gave Jirka a shake shouting that Hugo had gone mad and was about to drown. Jirka jumped up, he had been sleeping naked, and we rushed to the edge of the river. Jirka was about three meters away from Hugo when he turned to me and announced emphatically and reproachfully that this wasn't Hugo. At such a moment, a thousand things run through your head. The first thing that struck me was that this must be a criminal or a murderer. Jirka was closer to him than me and that was a good thing because Jirka was never afraid of anything or anyone. Just like my dad. I had seen him knock down much bigger lads to the ground when they had harmed him or harmed what he was fighting for, and he thumped them with his fists. And so it was a good thing that he was first. I was afraid. He moved closer to the bloke and the bloke, who was strong and really looked like Hugo, just waited. I noticed he had one hand under water. I began to be even more afraid and I suggested quietly that he might be holding an axe. That brought Jirka to a halt. They stood there in the river opposite each other for a long time. Jirka did not retreat, neither of them

evidently wanted to be a coward. And then the bloke took two or three steps back, gave a funny laugh again and swam off to the other side of the river. On the opposite bank he evidently disappeared in the bushes. We went to wake up the fishermen. There, where the Japanese fishing rods had stood, was now an empty spot.

APPENDIX 6

Abbreviations:

POS	position	
SYLL	word size (in syllables)	
FC	word class:	1 = content word category 2 = content+function word category 3 = function word category
SPEAK	speaker:	1 = AV 2 = LG 3 = OU 4 = PK
TASK	1 = reading 2 = paraphrasing 3 = spontaneous speech	

A number following any of these abbreviation indicates what group of phrases the model is fitted for; e.g., POS3 means factor position in 3-word intonation phrases.

STUDY 2

GENERAL LINEAR MODELS

MODELS WITH MAIN FACTORS, AND 2-WAY AND 3-WAY INTERACTIONS

2-word IP's

Tests of Between-Subjects Effects

Dependent Variable: AR2

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1736.637[b]	161	10.787	6.552	.000
Intercept	2304.217	1	2304.217	1399.692	.000
POS2	50.410	1	50.410	30.622	.000
SYLLS2	160.956	5	32.191	19.555	.000
FC2	6.995	2	3.497	2.124	.120
SPEAK2	4.244	3	1.415	.859	.462
TASK2	1.814	2	.907	.551	.577
POS2 * SYLLS2	4.466	5	.893	.543	.744
POS2 * FC2	1.290	2	.645	.392	.676
POS2 * SPEAK2	6.082	3	2.027	1.231	.297
POS2 * TASK2	5.789	2	2.894	1.758	.173
SYLLS2 * FC2	19.077	6	3.179	1.931	.074
SYLLS2 * SPEAK2	7.578	12	.632	.384	.969
SYLLS2 * TASK2	18.276	8	2.285	1.388	.199
FC2 * SPEAK2	2.470	6	.412	.250	.959
FC2 * TASK2	5.755	4	1.439	.874	.479
SPEAK2 * TASK2	19.592	6	3.265	1.984	.066
POS2 * SYLLS2 * FC2	17.013	4	4.253	2.584	.036
POS2 * SYLLS2 * SPEAK2	15.253	9	1.695	1.029	.415
POS2 * SYLLS2 * TASK2	15.705	6	2.618	1.590	.148
POS2 * FC2 * SPEAK2	14.176	6	2.363	1.435	.199
POS2 * FC2 * TASK2	5.902	3	1.967	1.195	.311
POS2 * SPEAK2 * TASK2	9.595	6	1.599	.971	.444
SYLLS2 * FC2 * SPEAK2	21.885	15	1.459	.886	.580
SYLLS2 * FC2 * TASK2	24.182	9	2.687	1.632	.103
SYLLS2 * SPEAK2 * TASK2	24.766	19	1.303	.792	.718
FC2 * SPEAK2 * TASK2	29.553	12	2.463	1.496	.121
Error	935.059	568	1.646		
Total	28028.533	730			
Corrected Total	2671.696	729			

Tests of Between-Subjects Effects

Dependent Variable: AR2

Source	Noncent. Parameter	Observed Power[a]
Corrected Model	1054.917	1.000
Intercept	1399.692	1.000
POS2	30.622	.998
SYLLS2	97.773	1.000
FC2	4.249	.218
SPEAK2	2.578	.090
TASK2	1.102	.044
POS2 * SYLLS2	2.713	.071
POS2 * FC2	.784	.032
POS2 * SPEAK2	3.694	.145
POS2 * TASK2	3.516	.170
SYLLS2 * FC2	11.588	.482
SYLLS2 * SPEAK2	4.604	.080
SYLLS2 * TASK2	11.102	.394
FC2 * SPEAK2	1.501	.033
FC2 * TASK2	3.496	.113
SPEAK2 * TASK2	11.901	.499
POS2 * SYLLS2 * FC2	10.334	.499
POS2 * SYLLS2 * SPEAK2	9.265	.282
POS2 * SYLLS2 * TASK2	9.540	.373
POS2 * FC2 * SPEAK2	8.611	.323
POS2 * FC2 * TASK2	3.585	.139
POS2 * SPEAK2 * TASK2	5.829	.182
SYLLS2 * FC2 * SPEAK2	13.294	.352
SYLLS2 * FC2 * TASK2	14.690	.539
SYLLS2 * SPEAK2 * TASK2	15.044	.364
FC2 * SPEAK2 * TASK2	17.952	.602
Error		
Total		
Corrected Total		

a. Computed using alpha = .01

b. R Squared = .650 (Adjusted R Squared = .551)

3-word IP's

Tests of Between-Subjects Effects

Dependent Variable: AR3

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	2331.346[b]	205	11.372	5.440	.000
Intercept	3659.757	1	3659.757	1750.512	.000
POS3	116.027	2	58.013	27.749	.000
SYLLS3	111.654	5	22.331	10.681	.000
FC3	49.619	2	24.810	11.867	.000
SPEAK3	22.875	3	7.625	3.647	.012
TASK3	.382	2	.191	.091	.913
POS3 * SYLLS3	13.435	6	2.239	1.071	.378
POS3 * FC3	9.172	4	2.293	1.097	.357
POS3 * SPEAK3	31.397	6	5.233	2.503	.021
POS3 * TASK3	5.304	4	1.326	.634	.638
SYLLS3 * FC3	9.588	5	1.918	.917	.469
SYLLS3 * SPEAK3	29.347	12	2.446	1.170	.301
SYLLS3 * TASK3	8.267	7	1.181	.565	.785
FC3 * SPEAK3	11.647	6	1.941	.929	.474
FC3 * TASK3	6.894	4	1.724	.824	.510
SPEAK3 * TASK3	19.171	6	3.195	1.528	.166
POS3 * SYLLS3 * FC3	14.292	9	1.588	.760	.654
POS3 * SYLLS3 * SPEAK3	21.640	17	1.273	.609	.887
POS3 * SYLLS3 * TASK3	31.769	10	3.177	1.520	.128
POS3 * FC3 * SPEAK3	19.671	12	1.639	.784	.667
POS3 * FC3 * TASK3	21.321	8	2.665	1.275	.253
POS3 * SPEAK3 * TASK3	16.412	12	1.368	.654	.796
SYLLS3 * FC3 * SPEAK3	33.960	15	2.264	1.083	.368
SYLLS3 * FC3 * TASK3	33.257	8	4.157	1.988	.045
SYLLS3 * SPEAK3 * TASK3	31.934	20	1.597	.764	.759
FC3 * SPEAK3 * TASK3	23.208	12	1.934	.925	.521
Error	1501.107	718	2.091		
Total	41640.386	924			
Corrected Total	3832.452	923			

Tests of Between-Subjects Effects

Dependent Variable: AR3

Source	Noncent. Parameter	Observed Power[a]
Corrected Model	1115.115	1.000
Intercept	1750.512	1.000
POS3	55.497	1.000
SYLLS3	53.406	1.000
FC3	23.734	.975
SPEAK3	10.941	.591
TASK3	.183	.014
POS3 * SYLLS3	6.426	.211
POS3 * FC3	4.387	.156
POS3 * SPEAK3	15.017	.651
POS3 * TASK3	2.537	.075
SYLLS3 * FC3	4.586	.144
SYLLS3 * SPEAK3	14.037	.440
SYLLS3 * TASK3	3.954	.094
FC3 * SPEAK3	5.571	.171
FC3 * TASK3	3.297	.105
SPEAK3 * TASK3	9.170	.354
POS3 * SYLLS3 * FC3	6.836	.178
POS3 * SYLLS3 * SPEAK3	10.351	.218
POS3 * SYLLS3 * TASK3	15.195	.538
POS3 * FC3 * SPEAK3	9.409	.240
POS3 * FC3 * TASK3	10.198	.351
POS3 * SPEAK3 * TASK3	7.850	.180
SYLLS3 * FC3 * SPEAK3	16.244	.476
SYLLS3 * FC3 * TASK3	15.907	.623
SYLLS3 * SPEAK3 * TASK3	15.274	.364
FC3 * SPEAK3 * TASK3	11.101	.311
Error		
Total		
Corrected Total		

a. Computed using alpha = .01

b. R Squared = .608 (Adjusted R Squared = .496)

4-word IP's

Tests of Between-Subjects Effects

Dependent Variable: AR4

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1300.511[b]	210	6.193	3.429	.000
Intercept	489.125	1	489.125	270.793	.000
POS4	25.153	3	8.384	4.642	.004
SYLLS4	29.953	5	5.991	3.317	.006
FC4	2.544	2	1.272	.704	.496
SPEAK4	3.244	3	1.081	.599	.616
TASK4	1.137	2	.568	.315	.730
POS4 * SYLLS4	27.373	10	2.737	1.515	.134
POS4 * FC4	1.343	6	.224	.124	.993
POS4 * SPEAK4	12.077	9	1.342	.743	.669
POS4 * TASK4	10.761	6	1.793	.993	.431
SYLLS4 * FC4	10.646	5	2.129	1.179	.320
SYLLS4 * SPEAK4	20.866	10	2.087	1.155	.322
SYLLS4 * TASK4	8.779	6	1.463	.810	.563
FC4 * SPEAK4	1.509	6	.252	.139	.991
FC4 * TASK4	3.920	4	.980	.543	.705
SPEAK4 * TASK4	5.002	6	.834	.462	.836
POS4 * SYLLS4 * FC4	31.406	6	5.234	2.898	.010
POS4 * SYLLS4 * SPEAK4	19.768	18	1.098	.608	.892
POS4 * SYLLS4 * TASK4	9.553	7	1.365	.756	.625
POS4 * FC4 * SPEAK4	12.478	15	.832	.461	.958
POS4 * FC4 * TASK4	10.875	6	1.812	1.003	.424
POS4 * SPEAK4 * TASK4	26.461	18	1.470	.814	.683
SYLLS4 * FC4 * SPEAK4	14.564	9	1.618	.896	.530
SYLLS4 * FC4 * TASK4	5.355	1	5.355	2.965	.086
SYLLS4 * SPEAK4 * TASK4	6.132	13	.472	.261	.996
FC4 * SPEAK4 * TASK4	34.079	11	3.098	1.715	.071
Error	428.086	237	1.806		
Total	20033.457	448			
Corrected Total	1728.597	447			

Tests of Between-Subjects Effects

Dependent Variable: AR4

Source	Noncent. Parameter	Observed Power[a]
Corrected Model	719.998	1.000
Intercept	270.793	1.000
POS4	13.925	.728
SYLLS4	16.583	.738
FC4	1.408	.056
SPEAK4	1.796	.058
TASK4	.629	.027
POS4 * SYLLS4	15.155	.514
POS4 * FC4	.743	.020
POS4 * SPEAK4	6.686	.164
POS4 * TASK4	5.957	.182
SYLLS4 * FC4	5.894	.201
SYLLS4 * SPEAK4	11.552	.353
SYLLS4 * TASK4	4.860	.135
FC4 * SPEAK4	.835	.021
FC4 * TASK4	2.170	.061
SPEAK4 * TASK4	2.769	.063
POS4 * SYLLS4 * FC4	17.387	.730
POS4 * SYLLS4 * SPEAK4	10.944	.212
POS4 * SYLLS4 * TASK4	5.289	.138
POS4 * FC4 * SPEAK4	6.908	.117
POS4 * FC4 * TASK4	6.021	.185
POS4 * SPEAK4 * TASK4	14.650	.340
SYLLS4 * FC4 * SPEAK4	8.063	.219
SYLLS4 * FC4 * TASK4	2.965	.193
SYLLS4 * SPEAK4 * TASK4	3.395	.049
FC4 * SPEAK4 * TASK4	18.867	.638
Error		
Total		
Corrected Total		

a. Computed using alpha = .01

b. R Squared = .752 (Adjusted R Squared = .533)

APPENDIX 7

STUDY 2

GENERAL LINEAR MODELS

MODELS WITH THE MAIN FACTORS AND SIGNIFICANT 2-WAY INTERACTIONS

2-word IP's

Tests of Between-Subjects Effects

Dependent Variable: AR2

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1432.965[a]	13	110.228	63.713	.000
Intercept	3247.136	1	3247.136	1876.880	.000
POS2	334.616	1	334.616	193.412	.000
SYLLS2	454.433	5	90.887	52.533	.000
FC2	119.388	2	59.694	34.504	.000
SPEAK2	33.714	3	11.238	6.496	.000
TASK2	6.811	2	3.406	1.969	.140
Error	1238.731	716	1.730		
Total	28028.533	730			
Corrected Total	2671.696	729			

a. R Squared = .536 (Adjusted R Squared = .528)

Tests of Between-Subjects Effects

Dependent Variable: AR2

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Corrected Model	1426.154[b]	11	129.650	74.738	.000	822.114	1.000
Intercept	3301.342	1	3301.342	1903.077	.000	1903.077	1.000
POS2	333.600	1	333.600	192.306	.000	192.306	1.000
SYLLS2	452.517	5	90.503	52.171	.000	260.856	1.000
FC2	124.658	2	62.329	35.930	.000	71.860	1.000
SPEAK2	33.726	3	11.242	6.481	.000	19.442	.971
Error	1245.543	718	1.735				
Total	28028.533	730					
Corrected Total	2671.696	729					

a. Computed using alpha = .05

b. R Squared = .534 (Adjusted R Squared = .527)

Post Hoc Tests

SYLLS2

Multiple Comparisons

Dependent Variable: AR2
Tukey HSD

(I) SYLLS2	(J) SYLLS2	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	-1.9851*	.162	.000	-2.4463	-1.5239
	3.00	-2.8571*	.166	.000	-3.3304	-2.3838
	4.00	-3.2160*	.192	.000	-3.7642	-2.6677
	5.00	-2.9469*	.359	.000	-3.9688	-1.9250
	6.00	-4.4205*	.774	.000	-6.6250	-2.2160
2.00	1.00	1.9851*	.162	.000	1.5239	2.4463
	3.00	-.8719*	.116	.000	-1.2023	-.5416
	4.00	-1.2308*	.151	.000	-1.6618	-.7999
	5.00	-.9618	.338	.051	-1.9258	2.262E-03
	6.00	-2.4354*	.764	.018	-4.6136	-.2571
3.00	1.00	2.8571*	.166	.000	2.3838	3.3304
	2.00	.8719*	.116	.000	.5416	1.2023
	4.00	-.3589	.156	.192	-.8027	8.492E-02
	5.00	-8.9836E-02	.340	1.000	-1.0597	.8800
	6.00	-1.5634	.765	.318	-3.7443	.6174
4.00	1.00	3.2160*	.192	.000	2.6677	3.7642
	2.00	1.2308*	.151	.000	.7999	1.6618
	3.00	.3589	.156	.192	-8.4919E-02	.8027
	5.00	.2691	.354	.974	-.7395	1.2776
	6.00	-1.2045	.771	.624	-3.4029	.9938
5.00	1.00	2.9469*	.359	.000	1.9250	3.9688
	2.00	.9618	.338	.051	-2.2619E-03	1.9258
	3.00	8.984E-02	.340	1.000	-.8800	1.0597
	4.00	-.2691	.354	.974	-1.2776	.7395
	6.00	-1.4736	.829	.480	-3.8350	.8878
6.00	1.00	4.4205*	.774	.000	2.2160	6.6250
	2.00	2.4354*	.764	.018	.2571	4.6136
	3.00	1.5634	.765	.318	-.6174	3.7443
	4.00	1.2045	.771	.624	-.9938	3.4029
	5.00	1.4736	.829	.480	-.8878	3.8350

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR2

Tukey HSD[a,b,c]

SYLLS2	N	Subset 1	Subset 2	Subset 3
1.00	86	3.6582		
2.00	288		5.6433	
3.00	234		6.5152	
5.00	16		6.6051	
4.00	103		6.8741	6.8741
6.00	3			8.0787
Sig.		1.000	.129	.146

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 1.735.

a. Uses Harmonic Mean Sample Size = 14.120.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

c. Alpha = .05.

FC2

Multiple Comparisons

Dependent Variable: AR2
Tukey HSD

(I) FC2	(J) FC2	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1.00	2.00	-1.5972*	.127	.000	-1.8942	-1.3002
	3.00	-1.5555*	.118	.000	-1.8315	-1.2795
2.00	1.00	1.5972*	.127	.000	1.3002	1.8942
	3.00	4.175E-02	.146	.956	-.2999	.3834
3.00	1.00	1.5555*	.118	.000	1.2795	1.8315
	2.00	-4.1745E-02	.146	.956	-.3834	.2999

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR2

Tukey HSD[a,b,c]

FC2	N	Subset 1	Subset 2
1.00	400	5.1820	
3.00	182		6.7375
2.00	148		6.7793
Sig.		1.000	.945

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 1.735.

a. Uses Harmonic Mean Sample Size = 203.372.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

c. Alpha = .05.

SPEAK2

Multiple Comparisons

Dependent Variable: AR2
Tukey HSD

(I) SPEAK2	(J) SPEAK2	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	-.1143	.142	.853	-.4803	.2516
	3.00	9.711E-02	.141	.902	-.2658	.4600
	4.00	-.4497*	.136	.005	-.7988	-.1006
2.00	1.00	.1143	.142	.853	-.2516	.4803
	3.00	.2114	.141	.437	-.1503	.5732
	4.00	-.3354	.135	.064	-.6834	1.257E-02
3.00	1.00	-9.7109E-02	.141	.902	-.4600	.2658
	2.00	-.2114	.141	.437	-.5732	.1503
	4.00	-.5468*	.134	.000	-.8916	-.2021
4.00	1.00	.4497*	.136	.005	.1006	.7988
	2.00	.3354	.135	.064	-1.2567E-02	.6834
	3.00	.5468*	.134	.000	.2021	.8916

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR2

Tukey HSD[a,b,c]

SPEAK2	N	Subset 1	Subset 2
3.00	178	5.6639	
1.00	170	5.7610	
2.00	172	5.8754	5.8754
4.00	210		6.2108
Sig.		.421	.073

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 1.735.

a. Uses Harmonic Mean Sample Size = 181.191.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

c. Alpha = .05.

3-word IP's

Tests of Between-Subjects Effects

Dependent Variable: AR3

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Corrected Model	1986.233[b]	20	99.312	48.574	.000	971.482	1.000
Intercept	2769.060	1	2769.060	1354.368	.000	1354.368	1.000
POS3	579.240	2	289.620	141.655	.000	283.311	1.000
SYLLS3	437.019	5	87.404	42.750	.000	213.749	1.000
FC3	276.919	2	138.459	67.721	.000	135.443	1.000
SPEAK3	83.831	3	27.944	13.667	.000	41.002	1.000
TASK3	16.676	2	8.338	4.078	.017	8.156	.725
POS3 * SPEAK3	59.698	6	9.950	4.866	.000	29.199	.992
Error	1846.219	903	2.045				
Total	41640.386	924					
Corrected Total	3832.452	923					

a. Computed using alpha = .05

b. R Squared = .518 (Adjusted R Squared = .508)

Post Hoc Tests

POS3

Multiple Comparisons

Dependent Variable: AR3
Tukey HSD

(I) POS3	(J) POS3	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	.4948*	.115	.000	.2247	.7648
	3.00	2.6626*	.115	.000	2.3926	2.9326
2.00	1.00	-.4948*	.115	.000	-.7648	-.2247
	3.00	2.1678*	.115	.000	1.8978	2.4379
3.00	1.00	-2.6626*	.115	.000	-2.9326	-2.3926
	2.00	-2.1678*	.115	.000	-2.4379	-1.8978

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR3

Tukey HSD[a,b]

POS3	N	Subset		
		1	2	3
3.00	308	4.7866		
2.00	308		6.9544	
1.00	308			7.4492
Sig.		1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 2.045.

a. Uses Harmonic Mean Sample Size = 308.000.

b. Alpha = .05.

SYLLS3

Multiple Comparisons

Dependent Variable: AR3
Tukey HSD

(I) SYLLS3	(J) SYLLS3	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	-2.3474*	.182	.000	-2.8660	-1.8288
	3.00	-2.6433*	.188	.000	-3.1794	-2.1071
	4.00	-3.1687*	.218	.000	-3.7910	-2.5465
	5.00	-3.4929*	.369	.000	-4.5438	-2.4419
	6.00	-4.3083*	1.025	.000	-7.2293	-1.3873
2.00	1.00	2.3474*	.182	.000	1.8288	2.8660
	3.00	-.2959	.108	.069	-.6044	1.267E-02
	4.00	-.8214*	.155	.000	-1.2629	-.3798
	5.00	-1.1455*	.335	.008	-2.1006	-.1904
	6.00	-1.9609	1.013	.381	-4.8488	.9270
3.00	1.00	2.6433*	.188	.000	2.1071	3.1794
	2.00	.2959	.108	.069	-1.2669E-02	.6044
	4.00	-.5255*	.162	.015	-.9875	-6.3428E-02
	5.00	-.8496	.339	.121	-1.8144	.1151
	6.00	-1.6650	1.015	.571	-4.5562	1.2261
4.00	1.00	3.1687*	.218	.000	2.5465	3.7910
	2.00	.8214*	.155	.000	.3798	1.2629
	3.00	.5255*	.162	.015	6.343E-02	.9875
	5.00	-.3241	.356	.944	-1.3393	.6910
	6.00	-1.1396	1.021	.875	-4.0479	1.7688
5.00	1.00	3.4929*	.369	.000	2.4419	4.5438
	2.00	1.1455*	.335	.008	.1904	2.1006
	3.00	.8496	.339	.121	-.1151	1.8144
	4.00	.3241	.356	.944	-.6910	1.3393
	6.00	-.8154	1.063	.973	-3.8445	2.2137
6.00	1.00	4.3083*	1.025	.000	1.3873	7.2293
	2.00	1.9609	1.013	.381	-.9270	4.8488
	3.00	1.6650	1.015	.571	-1.2261	4.5562
	4.00	1.1396	1.021	.875	-1.7688	4.0479
	5.00	.8154	1.063	.973	-2.2137	3.8445

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR3

Tukey HSD[a,b,c]

SYLLS3	N	Subset 1	Subset 2	Subset 3
1.00	72	4.0167		
2.00	433		6.3641	
3.00	292		6.6600	6.6600
4.00	106		7.1854	7.1854
5.00	19		7.5096	7.5096
6.00	2			8.3250
Sig.		1.000	.453	.087

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 2.045.

a. Uses Harmonic Mean Sample Size = 10.315.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

c. Alpha = .05.

FC3

Multiple Comparisons

Dependent Variable: AR3
Tukey HSD

(I) FC3	(J) FC3	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1.00	2.00	-1.1087*	.125	.000	-1.4023	-.8152
	3.00	-1.8841*	.111	.000	-2.1432	-1.6250
2.00	1.00	1.1087*	.125	.000	.8152	1.4023
	3.00	-.7754*	.140	.000	-1.1037	-.4470
3.00	1.00	1.8841*	.111	.000	1.6250	2.1432
	2.00	.7754*	.140	.000	.4470	1.1037

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR3

Tukey HSD[a,b,c]

FC3	N	Subset 1	Subset 2	Subset 3
1.00	494	5.6684		
2.00	177		6.7772	
3.00	253			7.5525
Sig.		1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 2.045.

a. Uses Harmonic Mean Sample Size = 258.029.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

c. Alpha = .05.

SPEAK3

Multiple Comparisons

Dependent Variable: AR3
Tukey HSD

(I) SPEAK3	(J) SPEAK3	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1.00	2.00	-.1365	.139	.762	-.4947	.2217
	3.00	.1865	.139	.539	-.1717	.5447
	4.00	-.5156*	.137	.001	-.8665	-.1648
2.00	1.00	.1365	.139	.762	-.2217	.4947
	3.00	.3230	.131	.067	-1.4466E-02	.6604
	4.00	-.3791*	.128	.017	-.7087	-4.9539E-02
3.00	1.00	-.1865	.139	.539	-.5447	.1717
	2.00	-.3230	.131	.067	-.6604	1.447E-02
	4.00	-.7021*	.128	.000	-1.0317	-.3725
4.00	1.00	.5156*	.137	.001	.1648	.8665
	2.00	.3791*	.128	.017	4.954E-02	.7087
	3.00	.7021*	.128	.000	.3725	1.0317

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR3

Tukey HSD[a,b,c]

SPEAK3	N	Subset 1	Subset 2
3.00	237	6.0774	
1.00	189	6.2639	
2.00	237	6.4004	
4.00	261		6.7795
Sig.		.075	1.000

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 2.045.

a. Uses Harmonic Mean Sample Size = 227.774.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

c. Alpha = .05.

TASK3

Multiple Comparisons

Dependent Variable: AR3
Tukey HSD

(I) TASK3	(J) TASK3	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1.00	2.00	5.447E-02	.121	.895	-.2301	.3390
	3.00	-.2421	.114	.086	-.5096	2.547E-02
2.00	1.00	-5.4471E-02	.121	.895	-.3390	.2301
	3.00	-.2965*	.113	.024	-.5616	-3.1470E-02
3.00	1.00	.2421	.114	.086	-2.5466E-02	.5096
	2.00	.2965*	.113	.024	3.147E-02	.5616

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR3

Tukey HSD[a,b,c]

TASK3	N	Subset 1	Subset 2
2.00	282	6.2622	
1.00	273	6.3166	6.3166
3.00	369		6.5587
Sig.		.886	.094

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 2.045.

a. Uses Harmonic Mean Sample Size = 302.446.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

c. Alpha = .05.

4-word IP's

Tests of Between-Subjects Effects

Dependent Variable: AR4

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Corrected Model	938.109[b]	14	67.008	36.893	.000	516.506	1.000
Intercept	4799.398	1	4799.398	2642.465	.000	2642.465	1.000
POS4	245.375	3	81.792	45.033	.000	135.099	1.000
SYLLS4	306.045	4	76.511	42.126	.000	168.503	1.000
FC4	108.727	2	54.363	29.932	.000	59.863	1.000
SPEAK4	56.391	3	18.797	10.349	.000	31.048	.999
TASK4	25.608	2	12.804	7.050	.001	14.099	.928
Error	784.624	432	1.816				
Total	19955.823	447					
Corrected Total	1722.733	446					

a. Computed using alpha = .05

b. R Squared = .545 (Adjusted R Squared = .530)

Post Hoc Tests

POS4

Multiple Comparisons

Dependent Variable: AR4
Tukey HSD

(I) POS4	(J) POS4	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	.4148	.180	.098	-4.8883E-02	.8785
	3.00	.7785*	.180	.000	.3148	1.2422
	4.00	2.6151*	.180	.000	2.1514	3.0788
2.00	1.00	-.4148	.180	.098	-.8785	4.888E-02
	3.00	.3637	.180	.181	-9.8975E-02	.8264
	4.00	2.2003*	.180	.000	1.7377	2.6630
3.00	1.00	-.7785*	.180	.000	-1.2422	-.3148
	2.00	-.3637	.180	.181	-.8264	9.898E-02
	4.00	1.8366*	.180	.000	1.3740	2.2993
4.00	1.00	-2.6151*	.180	.000	-3.0788	-2.1514
	2.00	-2.2003*	.180	.000	-2.6630	-1.7377
	3.00	-1.8366*	.180	.000	-2.2993	-1.3740

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR4

Tukey HSD[a,b,c]

POS4	N	Subset		
		1	2	3
4.00	112	4.7258		
3.00	112		6.5624	
2.00	112		6.9261	6.9261
1.00	111			7.3409
Sig.		1.000	.182	.098

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 1.816.

a. Uses Harmonic Mean Sample Size = 111.748.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

c. Alpha = .05.

SYLLS4

Multiple Comparisons

Dependent Variable: AR4
Tukey HSD

(I) SYLLS4	(J) SYLLS4	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	-2.5079*	.226	.000	-3.1253	-1.8905
	3.00	-2.9829*	.233	.000	-3.6174	-2.3484
	4.00	-3.4299*	.294	.000	-4.2323	-2.6275
	5.00	-4.2914*	.519	.000	-5.7069	-2.8759
2.00	1.00	2.5079*	.226	.000	1.8905	3.1253
	3.00	-.4750*	.144	.009	-.8690	-8.1033E-02
	4.00	-.9220*	.231	.001	-1.5517	-.2923
	5.00	-1.7835*	.486	.002	-3.1087	-.4583
3.00	1.00	2.9829*	.233	.000	2.3484	3.6174
	2.00	.4750*	.144	.009	8.103E-02	.8690
	4.00	-.4470	.237	.325	-1.0935	.1995
	5.00	-1.3085	.489	.057	-2.6418	2.479E-02
4.00	1.00	3.4299*	.294	.000	2.6275	4.2323
	2.00	.9220*	.231	.001	.2923	1.5517
	3.00	.4470	.237	.325	-.1995	1.0935
	5.00	-.8615	.521	.463	-2.2824	.5594
5.00	1.00	4.2914*	.519	.000	2.8759	5.7069
	2.00	1.7835*	.486	.002	.4583	3.1087
	3.00	1.3085	.489	.057	-2.4790E-02	2.6418
	4.00	.8615	.521	.463	-.5594	2.2824

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR4

Tukey HSD[a,b,c]

SYLLS4	N	Subset 1	Subset 2	Subset 3
1.00	43	3.8410		
2.00	202		6.3489	
3.00	153		6.8239	
4.00	41		7.2709	7.2709
5.00	8			8.1324
Sig.		1.000	.086	.128

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 1.816.

a. Uses Harmonic Mean Sample Size = 27.154.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

c. Alpha = .05.

FC4

Multiple Comparisons

Dependent Variable: AR4
Tukey HSD

(I) FC4	(J) FC4	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1.00	2.00	-1.2176*	.179	.000	-1.6379	-.7973
	3.00	-1.7146*	.152	.000	-2.0701	-1.3592
2.00	1.00	1.2176*	.179	.000	.7973	1.6379
	3.00	-.4971*	.203	.038	-.9733	-2.0763E-02
3.00	1.00	1.7146*	.152	.000	1.3592	2.0701
	2.00	.4971*	.203	.038	2.076E-02	.9733

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR4

Tukey HSD[a,b,c]

FC4	N	Subset 1	Subset 2	Subset 3
1.00	262	5.7571		
2.00	72		6.9747	
3.00	113			7.4718
Sig.		1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 1.816.

a. Uses Harmonic Mean Sample Size = 112.972.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

c. Alpha = .05.

SPEAK4

Multiple Comparisons

Dependent Variable: AR4
Tukey HSD

(I) SPEAK4	(J) SPEAK4	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1.00	2.00	.2055	.186	.685	-.2714	.6823
	3.00	.3534	.176	.185	-9.8881E-02	.8056
	4.00	-.4985*	.189	.042	-.9845	-1.2432E-02
2.00	1.00	-.2055	.186	.685	-.6823	.2714
	3.00	.1479	.174	.830	-.2983	.5941
	4.00	-.7039*	.187	.001	-1.1844	-.2235
3.00	1.00	-.3534	.176	.185	-.8056	9.888E-02
	2.00	-.1479	.174	.830	-.5941	.2983
	4.00	-.8518*	.178	.000	-1.3079	-.3958
4.00	1.00	.4985*	.189	.042	1.243E-02	.9845
	2.00	.7039*	.187	.001	.2235	1.1844
	3.00	.8518*	.178	.000	.3958	1.3079

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR4

Tukey HSD[a,b,c]

SPEAK4	N	Subset 1	Subset 2
3.00	136	6.0790	
2.00	108	6.2269	
1.00	103	6.4323	
4.00	100		6.9308
Sig.		.209	1.000

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 1.816.

a. Uses Harmonic Mean Sample Size = 110.129.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

c. Alpha = .05.

TASK4

Multiple Comparisons

Dependent Variable: AR4
Tukey HSD

(I) TASK4	(J) TASK4	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1.00	2.00	-.3648	.182	.111	-.7918	6.207E-02
	3.00	-.8050*	.157	.000	-1.1740	-.4360
2.00	1.00	.3648	.182	.111	-6.2071E-02	.7918
	3.00	-.4402*	.156	.013	-.8057	-7.4604E-02
3.00	1.00	.8050*	.157	.000	.4360	1.1740
	2.00	.4402*	.156	.013	7.460E-02	.8057

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR4

Tukey HSD[a,b,c]

TASK4	N	Subset 1	Subset 2
1.00	108	5.8855	
2.00	111	6.2503	
3.00	228		6.6905
Sig.		.071	1.000

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 1.816.

a. Uses Harmonic Mean Sample Size = 132.426.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

c. Alpha = .05.

APPENDIX 8

STUDY 2

GENERAL LINEAR MODELS

Individual speakers in 3-word phrases

Speaker AV

Tests of Between-Subjects Effects

Dependent Variable: AVAR3

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Corrected Model	408.308[b]	11	37.119	17.567	.000	193.241	1.000
Intercept	980.740	1	980.740	464.157	.000	464.157	1.000
AVPOS3	170.053	2	85.027	40.241	.000	80.481	1.000
AVSYLLS3	79.308	5	15.862	7.507	.000	37.534	.999
AVFC3	61.315	2	30.657	14.509	.000	29.019	.999
AVTASK3	12.951	2	6.476	3.065	.049	6.129	.586
Error	373.992	177	2.113				
Total	8197.924	189					
Corrected Total	782.300	188					

a. Computed using alpha = .05

b. R Squared = .522 (Adjusted R Squared = .492)

Post Hoc Tests

AVPOS3

Multiple Comparisons

Dependent Variable: AVAR3
Tukey HSD

(I) AVPOS3	(J) AVPOS3	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	.1287	.259	.873	-.4783	.7357
	3.00	2.6535*	.259	.000	2.0465	3.2605
2.00	1.00	-.1287	.259	.873	-.7357	.4783
	3.00	2.5248*	.259	.000	1.9178	3.1318
3.00	1.00	-2.6535*	.259	.000	-3.2605	-2.0465
	2.00	-2.5248*	.259	.000	-3.1318	-1.9178

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AVAR3

Tukey HSD[a,b]

AVPOS3	N	Subset 1	Subset 2
3.00	63	4.5377	
2.00	63		7.0626
1.00	63		7.1913
Sig.		1.000	.873

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 2.113.

a. Uses Harmonic Mean Sample Size = 63.000.

b. Alpha = .05.

Speaker LG

Tests of Between-Subjects Effects

Dependent Variable: LGAR3

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Corrected Model	428.976[b]	11	38.998	18.069	.000	198.762	1.000
Intercept	1171.578	1	1171.578	542.839	.000	542.839	1.000
LGPOS3	106.585	2	53.292	24.693	.000	49.385	1.000
LGSYLLS3	151.824	5	30.365	14.069	.000	70.346	1.000
LGFC3	75.058	2	37.529	17.389	.000	34.778	1.000
LGTASK3	14.177	2	7.089	3.284	.039	6.569	.620
Error	485.604	225	2.158				
Total	10623.189	237					
Corrected Total	914.581	236					

a. Computed using alpha = .05

b. R Squared = .469 (Adjusted R Squared = .443)

Post Hoc Tests

LGPOS3

Multiple Comparisons

Dependent Variable: LGAR3

Tukey HSD

(I) LGPOS3	(J) LGPOS3	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	4.639E-02	.234	.979	-.5014	.5942
	3.00	2.1266*	.234	.000	1.5788	2.6745
2.00	1.00	-4.6392E-02	.234	.979	-.5942	.5014
	3.00	2.0802*	.234	.000	1.5324	2.6281
3.00	1.00	-2.1266*	.234	.000	-2.6745	-1.5788
	2.00	-2.0802*	.234	.000	-2.6281	-1.5324

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

LGAR3

Tukey HSD[a,b]

LGPOS3	N	Subset 1	Subset 2
3.00	79	4.9981	
2.00	79		7.0783
1.00	79		7.1247
Sig.		1.000	.979

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 2.158.

a. Uses Harmonic Mean Sample Size = 79.000.

b. Alpha = .05.

Speaker OU

Tests of Between-Subjects Effects

Dependent Variable: OUAR3

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	321.663[b]	10	32.166	20.858	.000
Intercept	2815.166	1	2815.166	1825.476	.000
OUPOS3	60.493	2	30.246	19.613	.000
OUSYLLS3	79.473	4	19.868	12.883	.000
OUFC3	71.834	2	35.917	23.290	.000
OUTASK3	1.097	2	.548	.356	.701
Error	348.527	226	1.542		
Total	9423.664	237			
Corrected Total	670.189	236			

Tests of Between-Subjects Effects

Dependent Variable: OUAR3

Source	Noncent. Parameter	Observed Power[a]
Corrected Model	208.580	1.000
Intercept	1825.476	1.000
OUPOS3	39.226	1.000
OUSYLLS3	51.534	1.000
OUFC3	46.580	1.000
OUTASK3	.711	.107
Error		
Total		
Corrected Total		

a. Computed using alpha = .05

b. R Squared = .480 (Adjusted R Squared = .457)

Post Hoc Tests

OUPOS3

Multiple Comparisons

Dependent Variable: OUAR3
Tukey HSD

(I) OUPOS3	(J) OUPOS3	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	.5881*	.198	.008	.1250	1.0512
	3.00	2.1645*	.198	.000	1.7015	2.6276
2.00	1.00	-.5881*	.198	.008	-1.0512	-.1250
	3.00	1.5765*	.198	.000	1.1134	2.0396
3.00	1.00	-2.1645*	.198	.000	-2.6276	-1.7015
	2.00	-1.5765*	.198	.000	-2.0396	-1.1134

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

OUAR3

Tukey HSD[a,b]

OUPOS3	N	Subset		
		1	2	3
3.00	79	4.8304		
2.00	79		6.4068	
1.00	79			6.9949
Sig.		1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 1.542.

a. Uses Harmonic Mean Sample Size = 79.000.

b. Alpha = .05.

Speaker PK

Tests of Between-Subjects Effects

Dependent Variable: PKAR3

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	822.165[b]	10	82.217	35.594	.000
Intercept	3786.494	1	3786.494	1639.266	.000
PKPOS3	303.536	2	151.768	65.704	.000
PKSYLLS3	165.693	4	41.423	17.933	.000
PKFC3	75.727	2	37.864	16.392	.000
PKTASK3	13.510	2	6.755	2.924	.056
Error	577.468	250	2.310		
Total	13395.609	261			
Corrected Total	1399.633	260			

Tests of Between-Subjects Effects

Dependent Variable: PKAR3

Source	Noncent. Parameter	Observed Power[a]
Corrected Model	355.936	1.000
Intercept	1639.266	1.000
PKPOS3	131.408	1.000
PKSYLLS3	71.733	1.000
PKFC3	32.784	1.000
PKTASK3	5.849	.567
Error		
Total		
Corrected Total		

a. Computed using alpha = .05

b. R Squared = .587 (Adjusted R Squared = .571)

Post Hoc Tests

PKPOS3

Multiple Comparisons

Dependent Variable: PKAR3
Tukey HSD

(I) PKPOS3	(J) PKPOS3	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	1.0823*	.230	.000	.5423	1.6224
	3.00	3.6081*	.230	.000	3.0680	4.1482
2.00	1.00	-1.0823*	.230	.000	-1.6224	-.5423
	3.00	2.5258*	.230	.000	1.9857	3.0659
3.00	1.00	-3.6081*	.230	.000	-4.1482	-3.0680
	2.00	-2.5258*	.230	.000	-3.0659	-1.9857

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

PKAR3

Tukey HSD[a,b]

PKPOS3	N	Subset		
		1	2	3
3.00	87	4.7349		
2.00	87		7.2607	
1.00	87			8.3430
Sig.		1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 2.310.

a. Uses Harmonic Mean Sample Size = 87.000.

b. Alpha = .05.

APPENDIX 9

STUDY 2

GENERAL LINEAR MODELS

PATTERN ABC

Speaker AV

Tests of Between-Subjects Effects

Dependent Variable: AR3AV

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	209.651[b]	9	23.295	15.563	.000
Intercept	1641.388	1	1641.388	1096.620	.000
POS3AV	89.272	2	44.636	29.821	.000
SYLLS3AV	23.991	3	7.997	5.343	.002
FC3AV	9.499	2	4.749	3.173	.048
TASK3AV	2.009	2	1.005	.671	.514
Error	110.761	74	1.497		
Total	3619.161	84			
Corrected Total	320.412	83			

Tests of Between-Subjects Effects

Dependent Variable: AR3AV

Source	Noncent. Parameter	Observed Power[a]
Corrected Model	140.069	1.000
Intercept	1096.620	1.000
POS3AV	59.643	1.000
SYLLS3AV	16.029	.921
FC3AV	6.346	.591
TASK3AV	1.342	.159
Error		
Total		
Corrected Total		

a. Computed using alpha = .05

b. R Squared = .654 (Adjusted R Squared = .612)

Post Hoc Tests

POS3AV

Multiple Comparisons

Dependent Variable: AR3AV
Tukey HSD

(I) POS3AV	(J) POS3AV	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	1.3793*	.327	.000	.5972	2.1613
	3.00	3.5261*	.327	.000	2.7441	4.3082
2.00	1.00	-1.3793*	.327	.000	-2.1613	-.5972
	3.00	2.1469*	.327	.000	1.3648	2.9289
3.00	1.00	-3.5261*	.327	.000	-4.3082	-2.7441
	2.00	-2.1469*	.327	.000	-2.9289	-1.3648

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR3AV

Tukey HSD[a,b]

POS3AV	N	Subset		
		1	2	3
3.00	28	4.3756		
2.00	28		6.5225	
1.00	28			7.9018
Sig.		1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 1.497.

a. Uses Harmonic Mean Sample Size = 28.000.

b. Alpha = .05.

Speaker LG

Tests of Between-Subjects Effects

Dependent Variable: AR3LG

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	169.582[b]	9	18.842	26.567	.000
Intercept	1048.089	1	1048.089	1477.742	.000
POS3LG	67.926	2	33.963	47.886	.000
SYLLS3LG	10.141	3	3.380	4.766	.005
FC3LG	9.574	2	4.787	6.750	.002
TASK3LG	5.215	2	2.607	3.676	.031
Error	46.101	65	.709		
Total	3148.984	75			
Corrected Total	215.683	74			

Tests of Between-Subjects Effects

Dependent Variable: AR3LG

Source	Noncent. Parameter	Observed Power[a]
Corrected Model	239.101	1.000
Intercept	1477.742	1.000
POS3LG	95.772	1.000
SYLLS3LG	14.298	.883
FC3LG	13.499	.905
TASK3LG	7.352	.657
Error		
Total		
Corrected Total		

a. Computed using alpha = .05

b. R Squared = .786 (Adjusted R Squared = .757)

Post Hoc Tests

POS3LG

Multiple Comparisons

Dependent Variable: AR3LG
Tukey HSD

(I) POS3LG	(J) POS3LG	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	1.3497*	.238	.000	.7784	1.9211
	3.00	3.4139*	.238	.000	2.8425	3.9852
2.00	1.00	-1.3497*	.238	.000	-1.9211	-.7784
	3.00	2.0642*	.238	.000	1.4928	2.6355
3.00	1.00	-3.4139*	.238	.000	-3.9852	-2.8425
	2.00	-2.0642*	.238	.000	-2.6355	-1.4928

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR3LG

Tukey HSD[a,b]

POS3LG	N	Subset		
		1	2	3
3.00	25	4.4278		
2.00	25		6.4920	
1.00	25			7.8417
Sig.		1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = .709.

a. Uses Harmonic Mean Sample Size = 25.000.

b. Alpha = .05.

Speaker OU

Tests of Between-Subjects Effects

Dependent Variable: AR3OU

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	214.841[b]	10	21.484	21.257	.000
Intercept	1219.097	1	1219.097	1206.208	.000
POS3OU	92.573	2	46.287	45.797	.000
SYLLS3OU	30.005	4	7.501	7.422	.000
FC3OU	9.330	2	4.665	4.616	.012
TASK3OU	1.137	2	.569	.563	.571
Error	101.068	100	1.011		
Total	4541.964	111			
Corrected Total	315.909	110			

Tests of Between-Subjects Effects

Dependent Variable: AR3OU

Source	Noncent. Parameter	Observed Power[a]
Corrected Model	212.569	1.000
Intercept	1206.208	1.000
POS3OU	91.595	1.000
SYLLS3OU	29.687	.995
FC3OU	9.232	.769
TASK3OU	1.125	.141
Error		
Total		
Corrected Total		

a. Computed using alpha = .05

b. R Squared = .680 (Adjusted R Squared = .648)

Post Hoc Tests

POS3OU

Multiple Comparisons

Dependent Variable: AR3OU
Tukey HSD

(I) POS3OU	(J) POS3OU	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	1.7298*	.234	.000	1.1737	2.2859
	3.00	3.1149*	.234	.000	2.5589	3.6710
2.00	1.00	-1.7298*	.234	.000	-2.2859	-1.1737
	3.00	1.3851*	.234	.000	.8291	1.9412
3.00	1.00	-3.1149*	.234	.000	-3.6710	-2.5589
	2.00	-1.3851*	.234	.000	-1.9412	-.8291

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR3OU

Tukey HSD[a,b]

POS3OU	N	Subset		
		1	2	3
3.00	37	4.6703		
2.00	37		6.0554	
1.00	37			7.7852
Sig.		1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 1.011.

a. Uses Harmonic Mean Sample Size = 37.000.

b. Alpha = .05.

Speaker PK

Tests of Between-Subjects Effects

Dependent Variable: AR3PK

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Corrected Model	568.758[b]	10	56.876	29.020	.000	290.196	1.000
Intercept	1911.076	1	1911.076	975.083	.000	975.083	1.000
POS3PK	233.686	2	116.843	59.616	.000	119.233	1.000
SYLLS3PK	32.810	4	8.203	4.185	.003	16.741	.912
FC3PK	20.504	2	10.252	5.231	.007	10.462	.822
TASK3PK	12.965	2	6.482	3.307	.040	6.615	.616
Error	207.751	106	1.960				
Total	6357.372	117					
Corrected Total	776.509	116					

a. Computed using alpha = .05

b. R Squared = .732 (Adjusted R Squared = .707)

Post Hoc Tests

POS3PK

Multiple Comparisons

Dependent Variable: AR3PK

Tukey HSD

(I) POS3PK	(J) POS3PK	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	2.5834*	.317	.000	1.8298	3.3370
	3.00	5.1076*	.317	.000	4.3540	5.8612
2.00	1.00	-2.5834*	.317	.000	-3.3370	-1.8298
	3.00	2.5242*	.317	.000	1.7706	3.2778
3.00	1.00	-5.1076*	.317	.000	-5.8612	-4.3540
	2.00	-2.5242*	.317	.000	-3.2778	-1.7706

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR3PK

Tukey HSD[a,b]

POS3PK	N	Subset 1	Subset 2	Subset 3
3.00	39	4.3626		
2.00	39		6.8867	
1.00	39			9.4702
Sig.		1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 1.960.

a. Uses Harmonic Mean Sample Size = 39.000.

b. Alpha = .05.

PATTERN BAC

Speaker AV

Tests of Between-Subjects Effects

Dependent Variable: AR3AV

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Corrected Model	249.449[b]	11	22.677	13.829	.000	152.123	1.000
Intercept	627.869	1	627.869	382.897	.000	382.897	1.000
POS3AV	150.450	2	75.225	45.875	.000	91.750	1.000
SYLLS3AV	30.368	5	6.074	3.704	.005	18.520	.908
FC3AV	15.660	2	7.830	4.775	.012	9.550	.775
TASK3AV	7.177	2	3.589	2.188	.121	4.377	.430
Error	98.387	60	1.640				
Total	3262.217	72					
Corrected Total	347.836	71					

a. Computed using alpha = .05

b. R Squared = .717 (Adjusted R Squared = .665)

Post Hoc Tests

POS3AV

Multiple Comparisons

Dependent Variable: AR3AV
Tukey HSD

(I) POS3AV	(J) POS3AV	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	-1.8699*	.370	.000	-2.7583	-.9815
	3.00	2.2051*	.370	.000	1.3167	3.0935
2.00	1.00	1.8699*	.370	.000	.9815	2.7583
	3.00	4.0750*	.370	.000	3.1866	4.9634
3.00	1.00	-2.2051*	.370	.000	-3.0935	-1.3167
	2.00	-4.0750*	.370	.000	-4.9634	-3.1866

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR3AV

Tukey HSD[a,b]

POS3AV	N	Subset 1	Subset 2	Subset 3
3.00	24	4.2688		
1.00	24		6.4739	
2.00	24			8.3438
Sig.		1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 1.640.

a. Uses Harmonic Mean Sample Size = 24.000.

b. Alpha = .05.

Speaker LG

Tests of Between-Subjects Effects

Dependent Variable: AR3LG

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Corrected Model	255.218[b]	9	28.358	13.230	.000	119.068	1.000
Intercept	1398.697	1	1398.697	652.541	.000	652.541	1.000
POS3LG	124.501	2	62.251	29.042	.000	58.084	1.000
SYLLS3LG	20.105	3	6.702	3.127	.031	9.380	.706
FC3LG	16.774	2	8.387	3.913	.024	7.826	.689
TASK3LG	1.220	2	.610	.285	.753	.569	.093
Error	158.616	74	2.143				
Total	4052.114	84					
Corrected Total	413.834	83					

a. Computed using alpha = .05

b. R Squared = .617 (Adjusted R Squared = .570)

Post Hoc Tests

POS3LG

Multiple Comparisons

Dependent Variable: AR3LG
Tukey HSD

(I) POS3LG	(J) POS3LG	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	-1.5095*	.391	.001	-2.4453	-.5736
	3.00	2.4500*	.391	.000	1.5141	3.3858
2.00	1.00	1.5095*	.391	.001	.5736	2.4453
	3.00	3.9594*	.391	.000	3.0236	4.8953
3.00	1.00	-2.4500*	.391	.000	-3.3858	-1.5141
	2.00	-3.9594*	.391	.000	-4.8953	-3.0236

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR3LG

Tukey HSD[a,b]

POS3LG	N	Subset		
		1	2	3
3.00	28	4.4448		
1.00	28		6.8948	
2.00	28			8.4042
Sig.		1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 2.143.

a. Uses Harmonic Mean Sample Size = 28.000.

b. Alpha = .05.

Speaker OU

Tests of Between-Subjects Effects

Dependent Variable: AR3OU

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Corrected Model	145.470[b]	10	14.547	11.230	.000	112.303	1.000
Intercept	538.025	1	538.025	415.357	.000	415.357	1.000
POS3OU	47.910	2	23.955	18.493	.000	36.987	1.000
SYLLS3OU	19.029	4	4.757	3.673	.009	14.690	.858
FC3OU	6.538	2	3.269	2.524	.088	5.047	.489
TASK3OU	3.676	2	1.838	1.419	.249	2.838	.294
Error	86.787	67	1.295				
Total	3019.165	78					
Corrected Total	232.257	77					

a. Computed using alpha = .05

b. R Squared = .626 (Adjusted R Squared = .571)

Post Hoc Tests

POS3OU

Multiple Comparisons

Dependent Variable: AR3OU

Tukey HSD

(I) POS3OU	(J) POS3OU	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	-1.1021*	.316	.002	-1.8587	-.3455
	3.00	1.8834*	.316	.000	1.1268	2.6400
2.00	1.00	1.1021*	.316	.002	.3455	1.8587
	3.00	2.9855*	.316	.000	2.2289	3.7421
3.00	1.00	-1.8834*	.316	.000	-2.6400	-1.1268
	2.00	-2.9855*	.316	.000	-3.7421	-2.2289

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR3OU

Tukey HSD[a,b]

POS3OU	N	Subset 1	Subset 2	Subset 3
3.00	26	4.3545		
1.00	26		6.2378	
2.00	26			7.3400
Sig.		1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 1.295.

a. Uses Harmonic Mean Sample Size = 26.000.

b. Alpha = .05.

Speaker PK

Tests of Between-Subjects Effects

Dependent Variable: AR3PK

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Corrected Model	271.569[b]	10	27.157	18.249	.000	182.488	1.000
Intercept	681.954	1	681.954	458.256	.000	458.256	1.000
POS3PK	116.534	2	58.267	39.154	.000	78.308	1.000
SYLLS3PK	26.430	4	6.607	4.440	.003	17.760	.925
FC3PK	13.150	2	6.575	4.418	.015	8.837	.746
TASK3PK	.297	2	.149	.100	.905	.200	.065
Error	122.028	82	1.488				
Total	4461.181	93					
Corrected Total	393.597	92					

a. Computed using alpha = .05

b. R Squared = .690 (Adjusted R Squared = .652)

Post Hoc Tests

POS3PK

Multiple Comparisons

Dependent Variable: AR3PK
Tukey HSD

(I) POS3PK	(J) POS3PK	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	-1.1233*	.310	.001	-1.8629	-.3837
	3.00	2.6661*	.310	.000	1.9264	3.4057
2.00	1.00	1.1233*	.310	.001	.3837	1.8629
	3.00	3.7894*	.310	.000	3.0498	4.5290
3.00	1.00	-2.6661*	.310	.000	-3.4057	-1.9264
	2.00	-3.7894*	.310	.000	-4.5290	-3.0498

Based on observed means. The error term is Error.

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

AR3PK

Tukey HSD[a,b]

POS3PK	N	Subset		
		1	2	3
3.00	31	4.4616		
1.00	31		7.1277	
2.00	31			8.2510
Sig.		1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares
The error term is Mean Square(Error) = 1.488.

a. Uses Harmonic Mean Sample Size = 31.000.

b. Alpha = .05.

APPENDIX 10

MATHEMATICAL MODELLING OF ARTICULATION RATE

Model 1

Tests of Between-Subjects Effects

Dependent Variable: LOGN_AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	110.112[b]	5	22.022	342.608	.000
Intercept	146.947	1	146.947	2286.086	.000
CONT	8.032	1	8.032	124.961	.000
FUNC	.106	1	.106	1.649	.199
SYLLS	37.217	1	37.217	578.992	.000
POS	29.324	1	29.324	456.194	.000
NWORDS	16.694	1	16.694	259.708	.000
Error	134.729	2096	6.428E-02		
Total	6855.098	2102			
Corrected Total	244.841	2101			

Tests of Between-Subjects Effects

Dependent Variable: LOGN_AR

Source	Noncent. Parameter	Observed Power[a]
Corrected Model	1713.038	1.000
Intercept	2286.086	1.000
CONT	124.961	1.000
FUNC	1.649	.250
SYLLS	578.992	1.000
POS	456.194	1.000
NWORDS	259.708	1.000
Error		
Total		
Corrected Total		

a. Computed using alpha = .05

b. R Squared = .450 (Adjusted R Squared = .448)

Parameter Estimates

Dependent Variable	Parameter	B	Std. Error	t	Sig.
LOGN_AR	Intercept	1.228	.034	36.425	.000
	[CONT=.00]	.205	.018	11.179	.000
	[CONT=1.00]	0[b]	.	.	.
	[FUNC=.00]	2.163E-02	.017	1.284	.199
	[FUNC=1.00]	0[b]	.	.	.
	SYLLS	.162	.007	24.062	.000
	POS	-.165	.008	-21.359	.000
	NWORDS	.136	.008	16.115	.000

Parameter Estimates

Dependent Variable	Parameter	95% Confidence Interval		Noncent. Parameter	Observed Power[a]
		Lower Bound	Upper Bound		
LOGN_AR	Intercept	1.161	1.294	36.425	1.000
	[CONT=.00]	.169	.241	11.179	1.000
	[CONT=1.00]	.	.	.	.
	[FUNC=.00]	-1.140E-02	5.466E-02	1.284	.250
	[FUNC=1.00]	.	.	.	.
	SYLLS	.149	.175	24.062	1.000
	POS	-.180	-.150	21.359	1.000
	NWORDS	.119	.152	16.115	1.000

a. Computed using alpha = .05

b. This parameter is set to zero because it is redundant.

Regression

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.671[a]	.450	.448	.2535

a. Predictors: (Constant), FUNC, NWORDS, SYLLS, POS, CONT

b. Dependent Variable: LOGN_AR

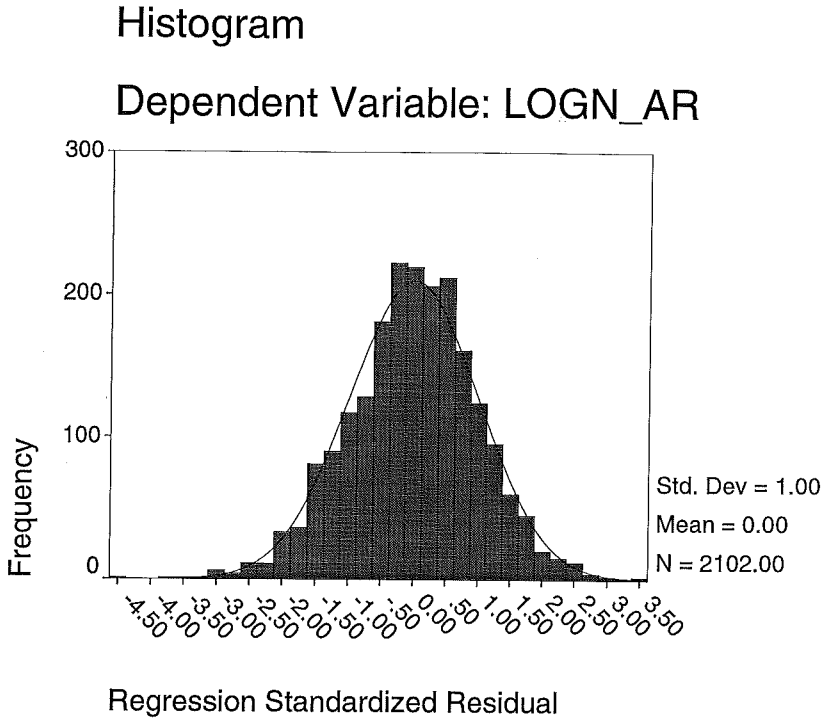
Residuals Statistics[a]

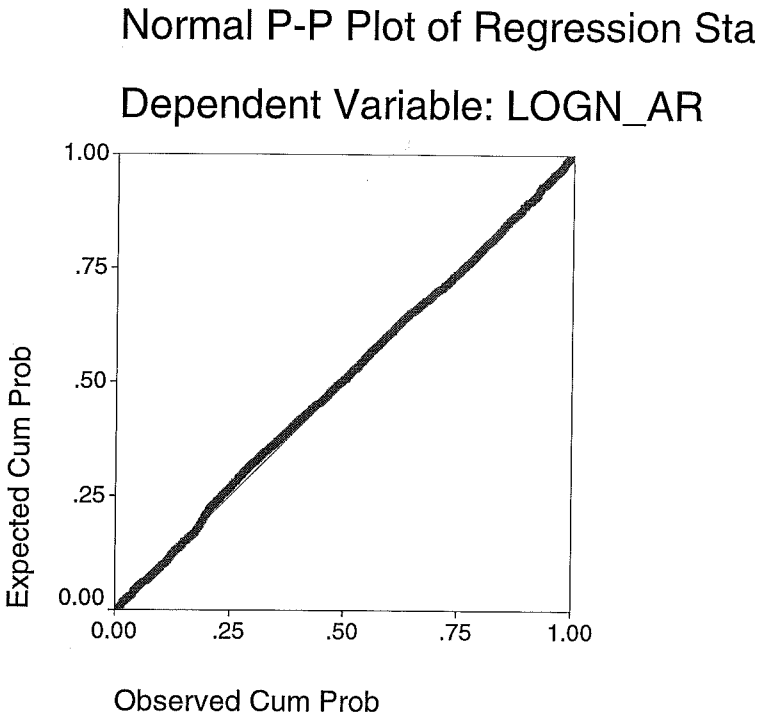
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	1.2929	2.6194	1.7733	.2289	2102
Residual	-1.1435	.9082	1.374E-15	.2532	2102
Std. Predicted Value	-2.098	3.696	.000	1.000	2102
Std. Residual	-4.510	3.582	.000	.999	2102

a. Dependent Variable: LOGN_AR

Oxford U. Speech
Phonetics Laboratory

Charts


Histogram
Dependent Variable: LOGN_AR
Frequency
300
200
100
0
Std. Dev = 1.00
Mean = 0.00
N = 2102.00
-4.50
-4.00
-3.50
-3.00
-2.50
-2.00
-1.50
-1.00
-.50
0.00
.50
1.00
1.50
2.00
2.50
3.00
3.50
Regression Standardized Residual


Normal P-P Plot of Regression Sta
Dependent Variable: LOGN_AR
Expected Cum Prob
1.00
.75
.50
.25
0.00
0.00
.25
.50
.75
1.00
Observed Cum Prob

Model 2

Tests of Between-Subjects Effects

Dependent Variable: LOGN_AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	102.080[b]	4	25.520	374.860	.000
Intercept	155.951	1	155.951	2290.748	.000
SYLLS	29.319	1	29.319	430.663	.000
POS	30.623	1	30.623	449.823	.000
NWORDS	17.535	1	17.535	257.564	.000
WCATBIN	4.369	1	4.369	64.179	.000
Error	142.761	2097	6.808E-02		
Total	6855.098	2102			
Corrected Total	244.841	2101			

Tests of Between-Subjects Effects

Dependent Variable: LOGN_AR

Source	Noncent. Parameter	Observed Power[a]
Corrected Model	1499.441	1.000
Intercept	2290.748	1.000
SYLLS	430.663	1.000
POS	449.823	1.000
NWORDS	257.564	1.000
WCATBIN	64.179	1.000
Error		
Total		
Corrected Total		

a. Computed using alpha = .05

b. R Squared = .417 (Adjusted R Squared = .416)

Parameter Estimates

Dependent Variable	Parameter	B	Std. Error	t	Sig.
LOGN_AR	Intercept	1.322	.029	45.073	.000
	SYLLS	.131	.006	20.752	.000
	POS	-.169	.008	-21.209	.000
	NWORDS	.139	.009	16.049	.000
	[WCATBIN=.00]	.104	.013	8.011	.000
	[WCATBIN=1.00]	0[b]	.	.	.

Parameter Estimates

Dependent Variable	Parameter	95% Confidence Interval		Noncent. Parameter	Observed Power[a]
		Lower Bound	Upper Bound		
LOGN_AR	Intercept	1.264	1.379	45.073	1.000
	SYLLS	.119	.144	20.752	1.000
	POS	-.184	-.153	21.209	1.000
	NWORDS	.122	.156	16.049	1.000
	[WCATBIN=.00]	7.829E-02	.129	8.011	1.000
	[WCATBIN=1.00]	.	.	.	.

a. Computed using alpha = .05

b. This parameter is set to zero because it is redundant.

Regression

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.646[a]	.417	.416	.2609

a. Predictors: (Constant), WCATBIN, NWORDS, SYLLS, POS

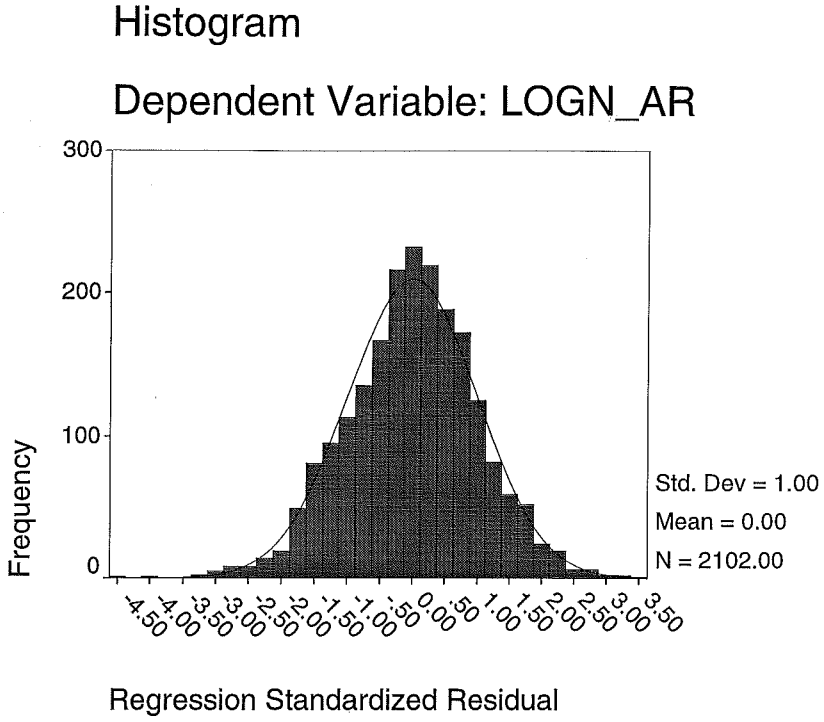
b. Dependent Variable: LOGN_AR

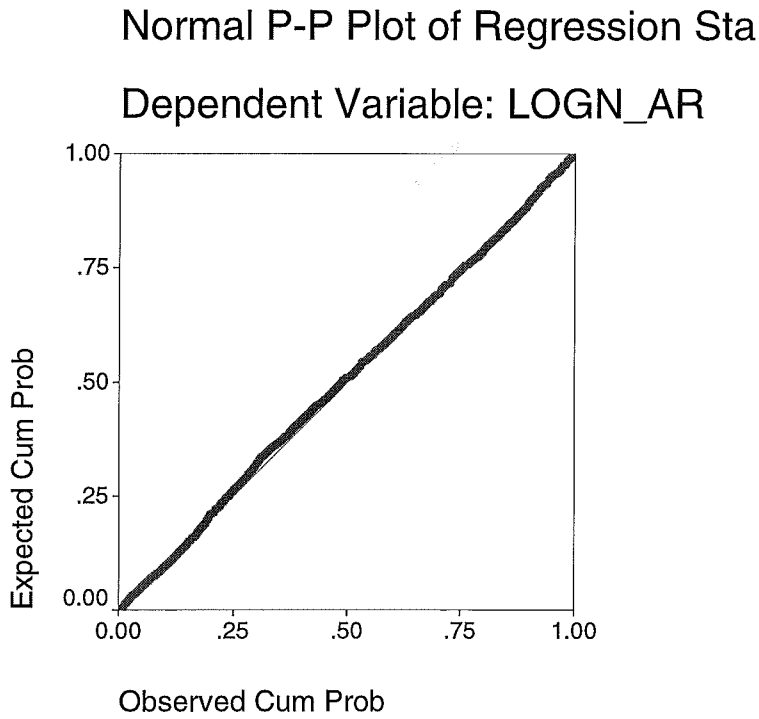
Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	1.3339	2.5992	1.7733	.2204	2102
Residual	-1.1847	.8944	1.450E-15	.2607	2102
Std. Predicted Value	-1.994	3.747	.000	1.000	2102
Std. Residual	-4.540	3.428	.000	.999	2102

a. Dependent Variable: LOGN_AR

Charts


Histogram
Dependent Variable: LOGN_AR
300
200
100
0
Frequency
Std. Dev = 1.00
Mean = 0.00
N = 2102.00
-4.50
-4.00
-3.50
-3.00
-2.50
-2.00
-1.50
-1.00
-.50
0.00
.50
1.00
1.50
2.00
2.50
3.00
3.50
Regression Standardized Residual


Normal P-P Plot of Regression Sta
Dependent Variable: LOGN_AR
1.00
.75
.50
.25
0.00
Expected Cum Prob
0.00
.25
.50
.75
1.00
Observed Cum Prob

Model 3

Tests of Between-Subjects Effects

Dependent Variable: LOGN_AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Corrected Model	117.304[b]	5	23.461	385.566	.000	1927.831	1.000
Intercept	17.061	1	17.061	280.389	.000	280.389	1.000
SYLL0.42	44.409	1	44.409	729.836	.000	729.836	1.000
POS	28.174	1	28.174	463.022	.000	463.022	1.000
CONT	8.959	1	8.959	147.232	.000	147.232	1.000
FUNC	.166	1	.166	2.726	.099	2.726	.379
NWORDS	15.913	1	15.913	261.513	.000	261.513	1.000
Error	127.537	2096	6.085E-02				
Total	6855.098	2102					
Corrected Total	244.841	2101					

a. Computed using alpha = .05

b. R Squared = .479 (Adjusted R Squared = .478)

Parameter Estimates

Dependent Variable	Parameter	B	Std. Error	t	Sig.
LOGN_AR	Intercept	.605	.050	12.173	.000
	SYLL0.42	.709	.026	27.015	.000
	POS	-.162	.008	-21.518	.000
	[CONT=.00]	.215	.018	12.134	.000
	[CONT=1.00]	0[b]	.	.	.
	[FUNC=.00]	2.688E-02	.016	1.651	.099
	[FUNC=1.00]	0[b]	.	.	.
	NWORDS	.132	.008	16.171	.000

Parameter Estimates

Dependent Variable	Parameter	95% Confidence Interval		Noncent. Parameter	Observed Power[a]
		Lower Bound	Upper Bound		
LOGN_AR	Intercept	.507	.702	12.173	1.000
	SYLL0.42	.658	.760	27.015	1.000
	POS	-.177	-.147	21.518	1.000
	[CONT=.00]	.180	.250	12.134	1.000
	[CONT=1.00]	.	.	.	.
	[FUNC=.00]	-5.047E-03	5.881E-02	1.651	.379
	[FUNC=1.00]	.	.	.	.
	NWORDS	.116	.148	16.171	1.000

a. Computed using alpha = .05

b. This parameter is set to zero because it is redundant.

Regression

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.692[a]	.479	.478	.2467

a. Predictors: (Constant), SYLL0.42, NWORDS, FUNC, POS, CONT

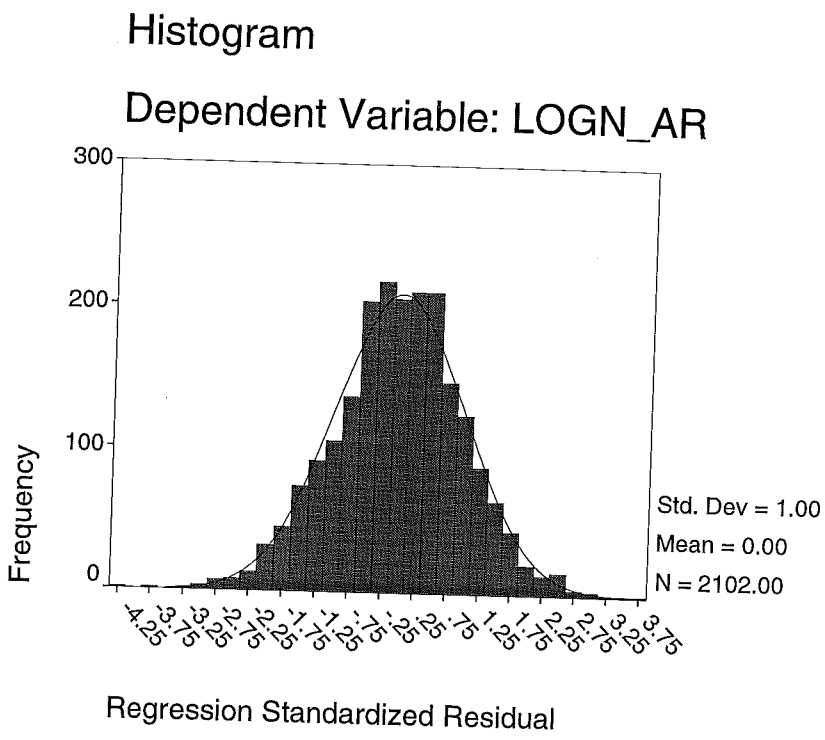
b. Dependent Variable: LOGN_AR

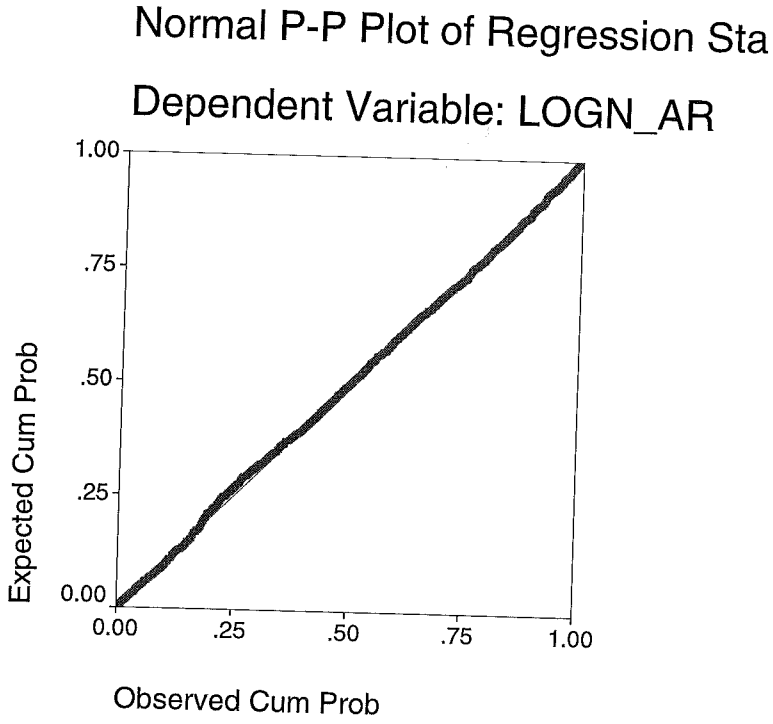
Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	1.2223	2.5812	1.7733	.2363	2102
Residual	-1.0729	.9044	-5.5309E-15	.2464	2102
Std. Predicted Value	-2.332	3.419	.000	1.000	2102
Std. Residual	-4.350	3.666	.000	.999	2102

a. Dependent Variable: LOGN_AR

Charts


Histogram
Dependent Variable: LOGN_AR
300
200
100
0
Frequency
Std. Dev = 1.00
Mean = 0.00
N = 2102.00
Regression Standardized Residual


Normal P-P Plot of Regression Sta
Dependent Variable: LOGN_AR
1.00
.75
.50
.25
0.00
Expected Cum Prob
0.00
.25
.50
.75
1.00
Observed Cum Prob

Model 4

Tests of Between-Subjects Effects

Dependent Variable: LOGN_AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	121.786[b]	5	24.357	414.877	.000
Intercept	181.551	1	181.551	3092.363	.000
POS	27.317	1	27.317	465.288	.000
FUNC	.137	1	.137	2.333	.127
NWORDS	15.264	1	15.264	259.998	.000
LNSYLL	48.891	1	48.891	832.759	.000
CONT	9.115	1	9.115	155.249	.000
Error	123.055	2096	5.871E-02		
Total	6855.098	2102			
Corrected Total	244.841	2101			

Tests of Between-Subjects Effects

Dependent Variable: LOGN_AR

Source	Noncent. Parameter	Observed Power[a]
Corrected Model	2074.387	1.000
Intercept	3092.363	1.000
POS	465.288	1.000
FUNC	2.333	.333
NWORDS	259.998	1.000
LNSYLL	832.759	1.000
CONT	155.249	1.000
Error		
Total		
Corrected Total		

a. Computed using alpha = .05

b. R Squared = .497 (Adjusted R Squared = .496)

Parameter Estimates

Dependent Variable	Parameter	B	Std. Error	t	Sig.
LOGN_AR	Intercept	1.266	.030	42.853	.000
	POS	-.160	.007	-21.571	.000
	[FUNC=.00]	2.425E-02	.016	1.527	.127
	[FUNC=1.00]	0[b]	.	.	.
	NWORDS	.130	.008	16.124	.000
	LNSYLL	.434	.015	28.858	.000
	[CONT=.00]	.215	.017	12.460	.000
	[CONT=1.00]	0[b]	.	.	.

Parameter Estimates

Dependent Variable	Parameter	95% Confidence Interval		Noncent. Parameter	Observed Power[a]
		Lower Bound	Upper Bound		
LOGN_AR	Intercept	1.208	1.324	42.853	1.000
	POS	-.174	-.145	21.571	1.000
	[FUNC=.00]	-6.886E-03	5.539E-02	1.527	.333
	[FUNC=1.00]	.	.	.	.
	NWORDS	.114	.145	16.124	1.000
	LNSYLL	.405	.464	28.858	1.000
	[CONT=.00]	.182	.249	12.460	1.000
	[CONT=1.00]	.	.	.	.

a. Computed using alpha = .05

b. This parameter is set to zero because it is redundant.

Regression

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.705[a]	.497	.496	.2423

a. Predictors: (Constant), LNSYLL, NWORDS, FUNC, POS, CONT

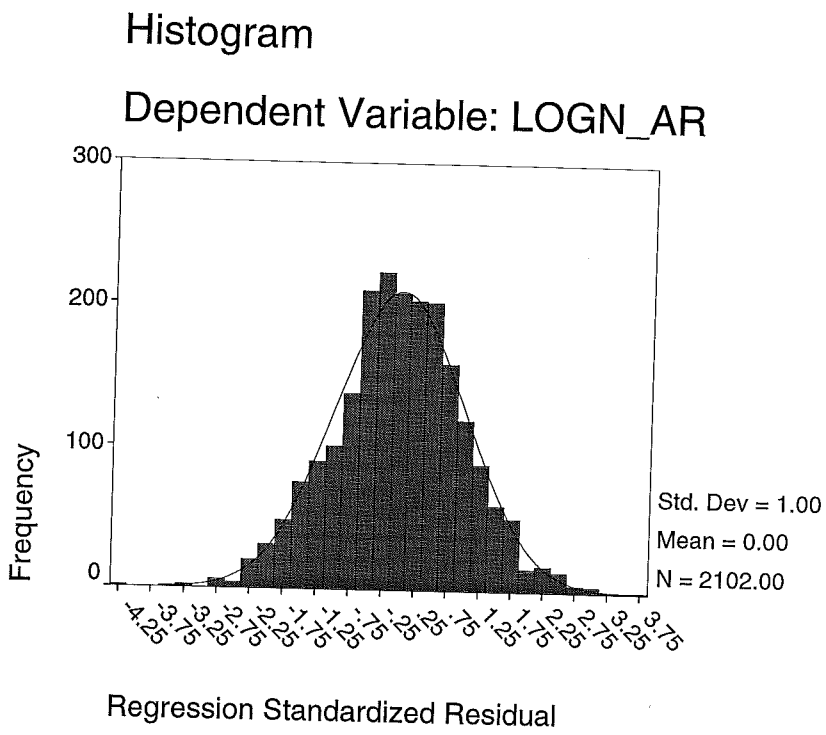
b. Dependent Variable: LOGN_AR

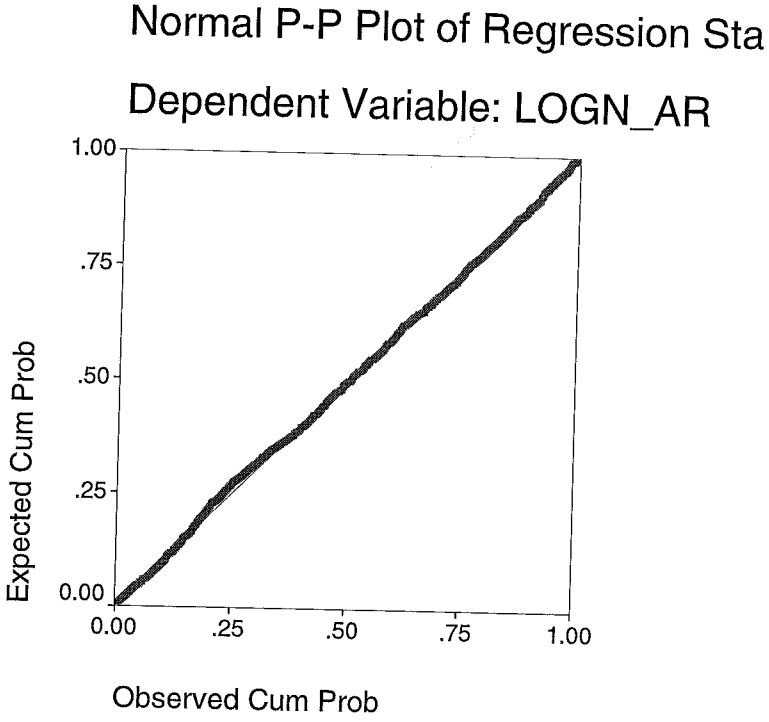
Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	1.1709	2.5397	1.7733	.2408	2102
Residual	-1.0218	.8965	-3.2692E-15	.2420	2102
Std. Predicted Value	-2.502	3.183	.000	1.000	2102
Std. Residual	-4.217	3.700	.000	.999	2102

a. Dependent Variable: LOGN_AR

Charts


Histogram
Dependent Variable: LOGN_AR
300
200
100
0
Frequency
Std. Dev = 1.00
Mean = 0.00
N = 2102.00
Regression Standardized Residual


Normal P-P Plot of Regression Sta
Dependent Variable: LOGN_AR
1.00
.75
.50
.25
0.00
Expected Cum Prob
0.00
.25
.50
.75
1.00
Observed Cum Prob

Model 5

Tests of Between-Subjects Effects

Dependent Variable: LOGN_AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Model	6726.665[b]	12	560.555	9121.952	.000
SYLLS	22.542	1	22.542	366.829	.000
POS	19.626	1	19.626	319.381	.000
WCAT	134.266	3	44.755	728.306	.000
NWORDS	10.400	1	10.400	169.246	.000
WCAT * POS	1.856	2	.928	15.099	.000
WCAT * NWORDS	3.047	2	1.524	24.795	.000
WCAT * SYLLS	2.904	2	1.452	23.625	.000
Error	128.433	2090	6.145E-02		
Total	6855.098	2102			

Tests of Between-Subjects Effects

Dependent Variable: LOGN_AR

Source	Noncent. Parameter	Observed Power[a]
Model	109463.430	1.000
SYLLS	366.829	1.000
POS	319.381	1.000
WCAT	2184.918	1.000
NWORDS	169.246	1.000
WCAT * POS	30.197	.999
WCAT * NWORDS	49.590	1.000
WCAT * SYLLS	47.250	1.000
Error		
Total		

a. Computed using alpha = .05

b. R Squared = .981 (Adjusted R Squared = .981)

Parameter Estimates

Dependent Variable	Parameter	B	Std. Error	t	Sig.
LOGN_AR	SYLLS	.176	.014	12.549	.000
	POS	-.148	.015	-9.736	.000
	[WCAT=1.00]	1.117	.037	30.583	.000
	[WCAT=2.00]	1.681	.073	22.891	.000
	[WCAT=3.00]	1.468	.055	26.937	.000
	NWORDS	.103	.015	6.671	.000
	[WCAT=1.00] * POS	-5.702E-02	.018	-3.104	.002
	[WCAT=2.00] * POS	4.875E-02	.023	2.116	.034
	[WCAT=3.00] * POS	0[b]	.	.	.
	[WCAT=1.00] * NWORDS	8.904E-02	.019	4.603	.000
	[WCAT=2.00] * NWORDS	-5.227E-02	.024	-2.193	.028
	[WCAT=3.00] * NWORDS	0[b]	.	.	.
	[WCAT=1.00] * SYLLS	1.053E-02	.016	.640	.522
	[WCAT=2.00] * SYLLS	-.109	.021	-5.222	.000
	[WCAT=3.00] * SYLLS	0[b]	.	.	.

Parameter Estimates

Dependent Variable	Parameter	95% Confidence Interval	
		Lower Bound	Upper Bound
LOGN_AR	SYLLS	.149	.204
	POS	-.177	-.118
	[WCAT=1.00]	1.046	1.189
	[WCAT=2.00]	1.537	1.825
	[WCAT=3.00]	1.361	1.575
	NWORDS	7.262E-02	.133
	[WCAT=1.00] * POS	-9.305E-02	-2.100E-02
	[WCAT=2.00] * POS	3.570E-03	9.394E-02
	[WCAT=3.00] * POS	.	.
	[WCAT=1.00] * NWORDS	5.111E-02	.127
	[WCAT=2.00] * NWORDS	-9.902E-02	-5.520E-03
	[WCAT=3.00] * NWORDS	.	.
	[WCAT=1.00] * SYLLS	-2.172E-02	4.278E-02
	[WCAT=2.00] * SYLLS	-.149	-6.777E-02
	[WCAT=3.00] * SYLLS	.	.

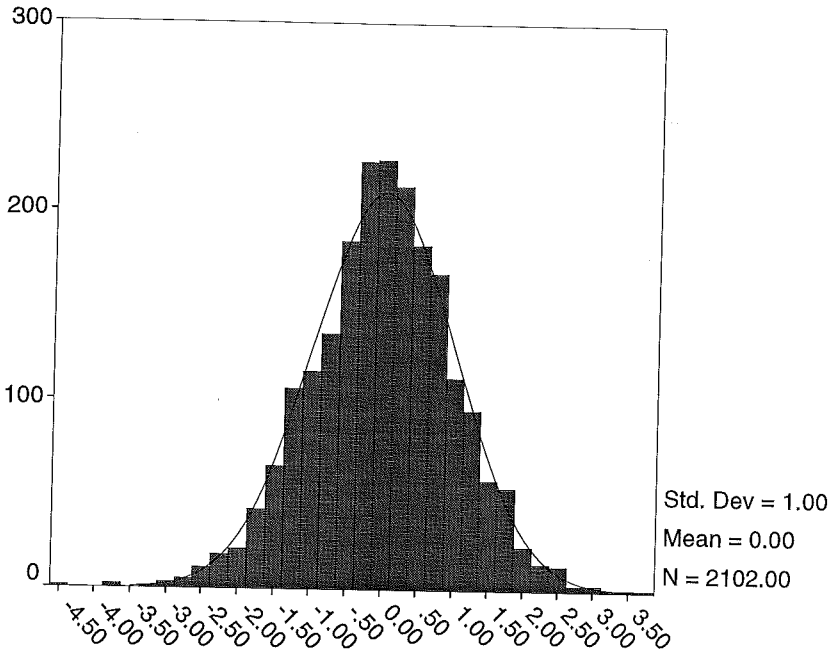
Parameter Estimates

Dependent Variable	Parameter	Noncent. Parameter	Observed Power[a]
LOGN_AR	SYLLS	12.549	1.000
	POS	9.736	1.000
	[WCAT=1.00]	30.583	1.000
	[WCAT=2.00]	22.891	1.000
	[WCAT=3.00]	26.937	1.000
	NWORDS	6.671	1.000
	[WCAT=1.00] * POS	3.104	.873
	[WCAT=2.00] * POS	2.116	.562
	[WCAT=3.00] * POS	.	.
	[WCAT=1.00] * NWORDS	4.603	.996
	[WCAT=2.00] * NWORDS	2.193	.592
	[WCAT=3.00] * NWORDS	.	.
	[WCAT=1.00] * SYLLS	.640	.098
	[WCAT=2.00] * SYLLS	5.222	.999
	[WCAT=3.00] * SYLLS	.	.

a. Computed using alpha = .05

b. This parameter is set to zero because it is redundant.

Graph



Standardized Residual for LOGN_AR

Model 6

Tests of Between-Subjects Effects

Dependent Variable: LOGN_AR

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Model	6736.695[b]	12	561.391	9909.446	.000
LNSYLL	20.207	1	20.207	356.681	.000
POS	18.516	1	18.516	326.844	.000
WCAT	164.516	3	54.839	967.988	.000
NWORDS	9.607	1	9.607	169.578	.000
WCAT * LNSYLL	1.452	2	.726	12.813	.000
WCAT * POS	1.698	2	.849	14.983	.000
WCAT * NWORDS	2.818	2	1.409	24.869	.000
Error	118.403	2090	5.665E-02		
Total	6855.098	2102			

Tests of Between-Subjects Effects

Dependent Variable: LOGN_AR

Source	Noncent. Parameter	Observed Power[a]
Model	118913.354	1.000
LNSYLL	356.681	1.000
POS	326.844	1.000
WCAT	2903.964	1.000
NWORDS	169.578	1.000
WCAT * LNSYLL	25.626	.997
WCAT * POS	29.965	.999
WCAT * NWORDS	49.739	1.000
Error		
Total		

a. Computed using alpha = .05

b. R Squared = .983 (Adjusted R Squared = .983)

Parameter Estimates

Dependent Variable	Parameter	B	Std. Error	t	Sig.
LOGN_AR	LNSYLL	.413	.028	14.810	.000
	POS	-.139	.015	-9.480	.000
	[WCAT=1.00]	1.186	.032	36.569	.000
	[WCAT=2.00]	1.649	.076	21.663	.000
	[WCAT=3.00]	1.564	.047	33.127	.000
	NWORDS	9.543E-02	.015	6.431	.000
	[WCAT=1.00] * LNSYLL	6.514E-02	.034	1.937	.053
	[WCAT=2.00] * LNSYLL	-.192	.056	-3.454	.001
	[WCAT=3.00] * LNSYLL	0[b]	.	.	.
	[WCAT=1.00] * POS	-6.086E-02	.018	-3.441	.001
	[WCAT=2.00] * POS	3.749E-02	.022	1.691	.091
	[WCAT=3.00] * POS	0[b]	.	.	.
	[WCAT=1.00] * NWORDS	8.977E-02	.019	4.827	.000
	[WCAT=2.00] * NWORDS	-4.378E-02	.023	-1.910	.056
	[WCAT=3.00] * NWORDS	0[b]	.	.	.

Parameter Estimates

Dependent Variable	Parameter	95% Confidence Interval	
		Lower Bound	Upper Bound
LOGN_AR	LNSYLL	.358	.467
	POS	-.167	-.110
	[WCAT=1.00]	1.122	1.250
	[WCAT=2.00]	1.499	1.798
	[WCAT=3.00]	1.472	1.657
	NWORDS	6.632E-02	.125
	[WCAT=1.00] * LNSYLL	-8.020E-04	.131
	[WCAT=2.00] * LNSYLL	-.302	-8.315E-02
	[WCAT=3.00] * LNSYLL	.	.
	[WCAT=1.00] * POS	-9.555E-02	-2.618E-02
	[WCAT=2.00] * POS	-5.987E-03	8.097E-02
	[WCAT=3.00] * POS	.	.
	[WCAT=1.00] * NWORDS	5.330E-02	.126
	[WCAT=2.00] * NWORDS	-8.872E-02	1.172E-03
	[WCAT=3.00] * NWORDS	.	.

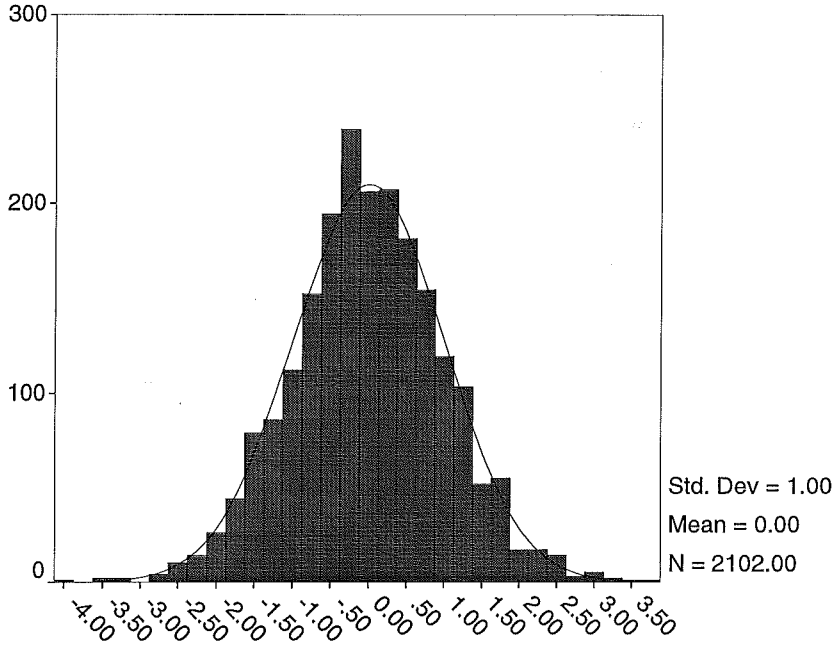
Parameter Estimates

Dependent Variable	Parameter	Noncent. Parameter	Observed Power[a]
LOGN_AR	LNSYLL	14.810	1.000
	POS	9.480	1.000
	[WCAT=1.00]	36.569	1.000
	[WCAT=2.00]	21.663	1.000
	[WCAT=3.00]	33.127	1.000
	NWORDS	6.431	1.000
	[WCAT=1.00] * LNSYLL	1.937	.491
	[WCAT=2.00] * LNSYLL	3.454	.932
	[WCAT=3.00] * LNSYLL	.	.
	[WCAT=1.00] * POS	3.441	.930
	[WCAT=2.00] * POS	1.691	.394
	[WCAT=3.00] * POS	.	.
	[WCAT=1.00] * NWORDS	4.827	.998
	[WCAT=2.00] * NWORDS	1.910	.480
	[WCAT=3.00] * NWORDS	.	.

a. Computed using alpha = .05

b. This parameter is set to zero because it is redundant.

Graph



Standardized Residual for LOGN_AR

REFERENCES

Abercrombie, D. (1967) *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.

Allen, G. D. (1975) Speech rhythm: Its relation to performance universals and articulatory timing, *Journal of Phonetics,* **3**, 75-86.

Amerman, J. D. & Parnell, M. (1992) Speech timing strategies in elderly adults, *Journal of Phonetics,* **20**, 65-76.

Arnfield, S., Roach, P., Setter, J., Greasley, P. & Horton, D. (1995) Emotional stress and speech tempo variation. In *Proceedings of ESCA-NATO Tutorial and Research Workshop on Speech under Stress*, pp. 13-15. Lisbon.

Bard, E. & Brew, C. (1990) *Psycholinguistic studies on incremental recognition of speech: An introduction to the messy and sticky.* University of Edinburgh.

Barik, H. C. (1977) Cross-linguistic study of temporal characteristics of different types of speech materials, *Language and Speech,* **20**, 116-126.

Bártková, K. (1991) Speaking rate modelization in French: Application to speech synthesis. In *Proceedings of 12th International Congress of Phonetic Sciences*, vol. 3, pp. 482-485. Aix-en-Provence.

Bartošek, J. (1974) *Jazyk rozhlasové sportovní publicistiky*. Praha.

Beattie, G. (1983) *Talk: An analysis of speech and non-verbal behaviour in conversation.* Milton Keynes: Open University Press.

Beattie, G.& Butterworth, B. (1979) Contextual probability and word frequency as determinants of pauses in spontaneous speech, *Language and Speech,* **22**, 201-211.

Beckman, M. & Pierrehumbert, J. (1986) Intonational structure in English and Japanese, *Phonology Yearbook,* **3**, 255-310.

Beckman, M. E. & Edwards, J. (1990) Lengthening and shortening and the nature of prosodic constituency. In *Papers in laboratory phonology I* (J. Kingston & M. E. Beckman, editors), pp. 152-178. Cambridge, UK: Cambridge University Press.

Bell-Berti, F. (1991) Comments on 'Some observation on the organisation an rhythm of speech. In *Proceedings of 12th International Congress of Phonetic Sciences,* vol. 2, pp. 238-242. Aix-en-Provence.

Bell-Berti, F., Gelfer, C. E. & Boyle, M. (1995) Utterance-final lengthening: the effect of speaking rate. In *Proceedings of 13th International Congress of Phonetic Sciences,* vol. 1, 162-165. Stockholm.

Berkovits, R. (1984) Duration and fundamental frequency in sentence-final intonation, *Journal of Phonetics,***12**, 255-265.

Berkovits, R. (1993) Utterance-final lengthening and the duration of final-stop closures, *Journal of Phonetics,* **21**, 479-489.

Berkovits, R. (1994) Durational effects in final lengthening, gapping and contrastive stress, *Language and Speech,* **37**, 237-250.

Birren, J. E., Woods, A. M. & Williams, M. V. (1979) Speed of behaviour as an indicator of age changes and the integrity of the nervous system. In *Brain function in old age* (F. Hofmeister & C. Mueller, editors). Berlin: Springer-Verlag.

Birren, J. E., Woods, A. M. & Williams, M. V. (1980) Behavioural slowing with age: causes, organization and consequences. In *Aging in the 1980's* (L. W. Poon, editor). Washington, D. C.: American Psychological Association.

Black, A. W. & Hunt, A. J. (1996) Generating F_0 contours from ToBI labels using linear regression. In *Proceedings of the 1996 International Conference on Spoken Language Processing (ICSLP)*, pp. 1385-1388. Philadelphia, Pennsylvania, USA.

Black, A. W. & Taylor, P. (1994) CHATR: a generic speech synthesis system. In *Proceedings of COLING-94*, vol.2, pp. 983-986. Kyoto, Japan.

Black, J. W. (1950) The effect of room characteristics upon vocal intensity and rate. *Journal of Acoustical Society of America,* **22**, 174-176.

Bolinger, D. L. (1963) Length, vowel, juncture, *Linguistics,* **1**, 5-29.

Bolinger, D. L. (1965) Pitch accent and sentence rhythm. In *Forms of English: Accent, Morpheme, Order* (I. Abe & T. Kanekiyo, editors), pp. 139-180. Cambridge, MA: Harvard University Press.

Bolinger, D. L. (1975) *Aspects of language.* London: Harcourt Brace Jovanovich.

Bolinger, D.L. (1981) *Two kinds of vowels, two kinds of rhythm.* Bloomington, Indiana: Indiana University Linguistics Club.

Boomer, D. S. (1965) Hesitation and grammatical encoding, *Language and Speech,* **8**, 148-158.

Boomer, D. S. (1970) Review of Goldman-Eisler (1968) Psycholinguistics: Experiments in spontaneous speech, *Lingua*, **25**, 152-164.

Boomer, D. S. & Laver, J. D. M. (1968/1973) Slips of the tongue. Reprinted in *Speech Errors as Linguistic Evidence.* (V. A. Fromkin, editor), pp. 120-131. The Hague: Mouton.

Brotherton, P. (1979) Speaking and not speaking: Processes for translation of ideas into speech. In *Of Speech and Time: Temporal Speech Patterns in Interpersonal Contexts* (A.W.Siegman & S. Feldstein, editors). Hillsdale, NJ: Lawrence Erlbaum.

Brown, G. (1977) *Listening to spoken English.* London: Longman.

Brubaker, R. S. (1972) Rate and pause characteristics of oral reading, *Journal of Psycholingustic Research,* **1**, 141-147.

Bruce, G., Granström, B. & House, D. (1991) Strategies for prosodic phrasing in Swedish. In *Proceedings of 12th International Congress of Phonetic Sciences,* vol. 4, pp. 182-185. Aix-en-Provence.

Butcher, A. (1981) Aspects of the speech pause: phonetic correlates and communicative function, *Arbeitsberichte,* **15**. Institut für Phonetik, Universität Kiel.

Butterworth, B. (1975). Hesitation and semantic planning in speech, *Journal of Psycholinguistic Research,* **4**, 75-87.

Butterworth, B. (1980a) Some constraints of models of language production. In *Language production: Vol. 1. Speech and talk* (B. Butterworth, editor), pp. 423-459. London: Academic Press.

Butterworth, B. (1980b) Evidence from pauses in speech. In *Language production, Vol.1: Speech and talk* (B. Butterworth, editor), pp. 155-176. London: Academic Press.

Buxton, H. (1983) Temporal predictability in the perception of English speech. In *Prosody: Models and Measurements* (A. Cutler & D. R. Ladd, editors), pp. 111-121. Heidelberg: Springer-Verlag.

Campbell, W. N. (1991) Durational shortening and anaphoric reference. In *Proceedings of 12th International Congress of Phonetic Sciences*, vol. 2, pp. 286-289. Aix-en-Provence.

Carlson, R. (1991) Duration models in use. In *Proceedings of 12th International Congress of Phonetic Sciences,* vol. 1, pp. 243-246. Aix-en-Provence.

Carlson, R., Granström, B. & Klatt, D. H. (1979) Some Notes on the Perception of Temporal Patterns in Speech. In *Frontiers of Speech Communication Research,* (B. Lindblom & S. Öhman, editors). London-New York-San Francisco: Academic Press.

Chafe, W. (1987) Cognitive constraints on information flow. In *Coherence and Grounding in Discourse,* (R. S. Tomlin, editor), pp. 21-51. Amsterdam: Benjamin.

Chafe, W., Du Bois, J. W. & Thompson, S. A. (1991) Towards a new corpus of spoken American English. In *English Corpus Linguistics* (K. Aijmer & B. Altenberg, editors). London: Longman.

Chan, D. *et al.* (1995) EUROM: A spoken language resource for the EU. In *Proceedings of Eurospeech 95,* pp. 867-870. Madrid.

Chlumsk″, J. (1911) Pokus o měření českých zvuk˘ a slabik v řeči souvislé. Praha.

Chlumsk″, J. (1928) Česká kvantita, melodie a přízvuk. Praha.

Church, K. W. (1983) *Phrase Structure Parsing: A Method for Taking Advantage of Allophonic Constraints.* PhD thesis, MIT.

Collier, R. (1975) Physiological correlates of intonation patterns. *Journal of Acoustical Society of America,* **58**, 249-255.

Collier, R. (1983) Physiological explanations of F0 declination. In *Abstracts of the 10th International Congress of Phonetic Sciences* (A. Cohen & M. P. R. van den Broecke, editors). Dordrecht: Foris. 440ff.

Cooper, W. E. & Paccia-Cooper, J. (1980) *Syntax and speech.* Cambridge, MA: Harvard University Press.

Cooper, W. E., Paccia, J. M. & Lapointe, S. G. (1978) Hierarchical coding in speech timing, *Cognitive Psychology,* **10**, 154-177.

Couper-Kuhlen (1986) *An introduction to English prosody.* London: Edward Arnold.

Crompton, A. (1980) Timing patterns in French, *Phonetica,* **37**, 205-234.

Crompton, A. (1981) Syllables and segments in speech production. *Linguistics*, **19-7/8**, 663-716.

Cruttenden, A. (1986) *Intonation.* Cambridge, UK: Cambridge University Press.

Crystal, D. & Quirk, R. (1964) *Systems of prosodic and paralinguistic features in English.* The Hague: Mouton.

Crystal, D. (1969) *Prosodic systems and intonation in English.* Cambridge, UK: Cambridge University Press.
Crystal, D. (1980) *A first dictionary of linguistics and phonetics.* Cambridge, UK: Cambridge University Press.
Crystal, D. (1985) *A Dictionary of Linguistics and Phonetics.* Oxford: Blackwell.
Crystal, T. H. & House, A. S. (1982) Segmental durations in connected speech signals: preliminary results. *Journal of the Acoustical Society of America,* **72**, 705-716.
Crystal, T. H. & House, A. S. (1988) Segmental durations in connected-speech signals: Current results. *Journal of Acoustical Society of America,* **83**(4), 1553-1573.
Crystal, T. H. & House, A. S. (1990) Articulation rate and the duration of syllables and stress groups in connected speech, *Journal of the Acoustical Society of America,* **88**(1), 101-112.
Dalton, P. & Hardcastle, W. J. (1977) *Disorders of fluency.* London: Edward Arnold.
Daneš, F. (1957) Intonace a věta ve spisovné češtině. Praha.
Dankovičová, J. (1992) The minimum pause duration in spontaneous speech, *PROPH - Progress Reports from Oxford Phonetics*, **5**, 17-24.
Dankovičová, J. (1994) Variability in articulation rate in spontaneous Czech speech. M. Phil. thesis. University of Oxford.
Dankovičová, J. (1995) The domain of articulation rate variability in Czech. In *Proceedings of 13th International Congress of Phonetic Sciences,* vol. 1, pp. 346-349. Stockholm.
Dankovičová, J. (1997) The domain of articulation rate variation in Czech, *Journal of Phonetics*, **25**, 287-312.
Dauer, R. M. (1983) Stress-timing and syllable-timing reanalysed, *Journal of Phonetics*, **11**, 51-69.
Dawson, L. O. (1929) A study of the development of the rate of articulation, *Elementary School Journal*, **29**, 610-615.
de Johnson, T., O'Connell, D. C. & Sabin, E. J. (1979) Temporal analysis of English and Spanish narratives. *Bulletin of the Psychonomic Society*, **13**, 347-350.
de Pijper, J. R. & Sanderman, A. (1995) On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues, *Journal of the Acoustical Society of America,* **96**, 2037-2047.
Deese, J. E. (1984) *Thought into Speech.* New Jersey: Prentice-Hall.
den Os, E. (1988) *Rhythm and tempo of Dutch and Italian.* Drukkerij Elinkwijk BV: Utrecht.
Duez, D. (1982) Silent and non-silent pauses in three speech styles, *Language and Speech*, **25**, 11-28.
Edwards, J. & Beckman, M. (1988). Articulatory timing and prosodic interpretation of syllabic duration, *Phonetica,* **45**, 156-174.
Eefting, W. (1988) Temporal variation in natural speech: Some explorations. In *Proc. Speech '88: 7th FASE Symposium*, pp. 503-507. Edinburgh.

Eefting, W. (1990) Production and perception of temporal variation: An explorative study. In *OTS Yearbook 1990,* pp. 13-38. Research Institute for Language and Speech: University of Utrecht.

Eefting, W. (1991) The effect of 'information value' and 'accentuation' on the duration of Dutch words, syllables, and segments, *Journal of the Acoustical Society of America,* **89**(1), 412-424.

Eefting, W. & Nooteboom, S. (1991) The effect of information value and accentuation on the word durations: A production and a perception study. In *Proceeding Proceedings of 12th International Congress of Phonetic Sciences*, vol. 3, pp. 302-305. Aix-en-Provence.

Eefting, W. & Rietveld, A. C. M. (1988) Just noticeable differences of articulation rate at sentence level. In *Proceedings*, **12**, Department of Language and Speech, Phonetics section, University of Nijmegen.

Fant, G. (1991) Units of temporal organization, stress groups versus syllables and words. In *Proceedings of 12th International Congress of Phonetic Sciences*, vol. 1, pp. 247-250. Aix-en-Provence.

Fant, G., Kruckenberg, A. & Nord, L. (1989) Preliminaries to the study of Swedish prose reading and reading style. *Speech Transmission Laboratory, Quarterly Progress and Status Report, Royal Institute of Technology (KTH), Stockholm,* **2**, 1-83.

Fant, G., Kruckenberg, A. & Nord, L. (1991a) Temporal organisation and rhythm in Swedish. In *Proceedings of 12th International Congress of Phonetic Sciences,* vol. 1, pp. 251-256. Aix-en-Provence.

Fant, G., Kruckenberg, A. & Nord, L.. (1991b) Language specific patterns of prosodic and segmental structures in Swedish, French and English. In *Proceedings of 12th International Congress of Phonetic Sciences*, vol. 4, pp. 118-121. Aix-en-Provence.

Fant, G, Kruckenberg, A & Nord, L. (1992) Prediction of syllable duration, speech rate and tempo. In *Proceedings of the 1992 International Conference on Spoken Language Processing (ICSLP),* vol. 1, pp. 667-670. University of Alberta, Canada.

Farnetani, E. (1989) Acoustic correlates of linguistic boundaries in Italian: a study on duration and fundamental frequency, *Eurospeech,* vol. 2, pp. 332-335. Paris.

Ferguson, G. A. & Takane, Y. (1989) *Statistical analysis in psychology and education.* Singapore: McGraw Hill Book Co.

Ferreira, F. (1993) Creation of prosody during sentence production, *Psychological Review,* **100**, 233-253.

Finegan, E. (1994) *Language: Its structure and use.* 2nd edition. USA: Harcourt Brace College Publishers.

Fletcher, J. (1991) Rhythm and final lengthening in French, *Journal of Phonetics,* **19**, 193-212.

Fletcher, S. (1972) Time-by-count measurement of diadochokinetic syllable rate, *Journal of Speech and Hearing Research*, **15**, 763-770.

Fónagy, I. & Magdics, K. (1960) Speed of utterance in phrases of different lengths, *Language and Speech*, **3**, 179-192.

Ford, M. & Holmes, V. M. (1978) Planning units in sentence production, *Cognition*, **6**, 35-53.
Fowler, C. (1988) Differential shortening of repeated content words produced in various communicative contexts, *Language and Speech,* **31**, 307-320.
Fowler, C. A. & Housum, J. (1987) Talkers' signalling of 'new' and 'old' words in speech and listeners' perception and use of the distinction, *Journal of Memory and Language*, **26**, 489-504.
Fowler, C. A. & Levy, E. T. (1991) Some ways in which forms arise from function in linguistic communications. In *Proceeding Proceedings of 12th International Congress of Phonetic Sciences*, vol. 2, pp. 279-282. Aix-en-Provence.
Fromkin, V. A. (1973) *Speech Errors as Linguistic Evidence*. The Hague: Mouton.
Garman, M. (1990) *Psycholinguistics*. Cambridge, UK: Cambridge University Press.
Garrett, M. F. (1980) Levels of processing in sentence production. In *Language Production, vol.1: speech and talk* (B. Butterworth, editor) London: Academic Press.
Garrido J.M., Llisterri, J., Marín, R., de la Mota, C. & Ríos, A. (1995) Prosodic markers at syntactic boundaries in Spanish. In *Proceedings of 13th International Congress of Phonetic Sciences,* vol. 2, pp. 370-373. Stockholm.
Gay, T. (1981) Mechanisms in the control of speech rate, *Phonetica*, **38**, 148-158.
Goldman-Eisler, F. (1954) On the variability of speed of talking and on its relation to the length of utterances in conversation, *British Journal of Psychology*, **45**, 45-94.
Goldman-Eisler, F. (1956) The determinants of the rate of speech output and their mutual relations, *Journal of Psychosomatic Research*, **1**, 137-143.
Goldman-Eisler, F. (1961a) The significance of changes in the rate of articulation, *Language and Speech*, **4**, 171-174.
Goldman-Eisler, F. (1961b) The continuity of speech utterance, its determinants and its significance, *Language and Speech*, **4**, 220-231.
Goldman-Eisler, F. (1967) Sequential temporal patterns and cognitive processes in speech, *Quarterly Journal of Experimental Psychology*, **10**, 96-106.
Goldman-Eisler, F. (1968) *Psycholinguistics: Experiments in spontaneous speech.* London and New York: Academic Press.
Goldman-Eisler, F. (1972) Pauses, clauses, sentences, *Language and Speech,* **15**, 103-113.
Grosjean, F. (1979) A study of timing in manual and a spoken language: American Sign Language and English, *Journal of Psycholinguistic Research,* **8**, 379-405.
Grosjean, F. (1980) Linguistic structures and performance structures: Studies in pause distribution. In *Temporal variables in speech: Studies in honour of F. Goldman-Eisler* (H. W. Dechert & M. Raupach, editors). The Hague: Mouton Publishers.
Grosjean, F. & Collins, M. (1979) Breathing, pausing and reading, *Phonetica*, **36**, 98-114.

Grosjean, F. & Deschamps, A. (1972) Analyse des variables temporelles du français spontané, *Phonetica*, **26**, 129-156.
Grosjean, F. & Deschamps, A. (1975) Analyse contrastive des variables temporelles de l'anglais et du français: vitesse de parole et variables composantes, phénomènes d'hesitation, *Phonetica*, **31**, 144-184.
Grosjean, F. & Gee, J. P. (1987) Prosodic structure and spoken word recognition. In *Spoken word recognition* (U. H. Frauenfelder & L. K. Tyler, editors), pp. 135-156. MIT Press.
Grosjean, F. & Lane, H. (1977) Pauses and syntax in American Sign Language, *Cognition,* **5**, 101-117.
Gussenhoven, C. & Rietveld, A. C. M. (1992) Intonation contours, prosodic structure and preboundary lengthening, *Journal of Phonetics,* **20**, 283-303.
Hála, B. (1962). Uvedení do fonetiky češtiny na obecně fonetickém základě. Praha.
Hallé, P., de Boysson-Bardies, B. & Vihman, M. M. (1991) Beginnings of prosodic organization: Intonation and duration patterns of disyllables produced by Japanese and French infants, *Language and Speech,* **34**, 299-318.
Halliday, M. A. K. (1967) Notes on transitivity and theme in English: II, *Journal of Linguistics*, **3**, 199-244.
Halliday, M. A. K. (1970) Language structure and language function. In *New Horizons in Linguistics* (J. Lyons, editor), pp. 140-165. Penguin.
Haselager, G. J. T., Slis, I. H. & Rietveld, A. C. M. (1991) An alternative method of studying the development of speech rate, *Clinical Linguistics and Phonetics*, **5**, 53-63.
Hawkins, S. & Warren, P. (1994) Phonetic influences on the intelligibility of conversational speech, *Journal of Phonetics,* **22**, 493-511.
Heike, A. E., Kowal, S. & O'Connell, D. C. (1983) The trouble with "articulatory" pauses, *Language and Speech*, **26**, 203-214.
Henze, R. (1951) Experimentelle Untersuchengen zur Phänomenologie der Sprechgeschwindigkeit. PhD thesis, Universität Göttingen.
Hirst, D. (1991) Intonation models. In *Proceedings of 12th International Congress of Phonetic Sciences*, vol. 1, pp. 305-310. Aix-en-Provence.
Hofhuis, E. (1993) Establishing prosodic structure by measuring segment duration. *Working papers,* vol. 42, pp. 136-139. Lund University, Department of Linguistics.
Hofhuis, E., Gussenhoven, C. & Rietveld, A. (1995) Final lengthening at prosodic boundaries in Dutch. In *Proceedings of 13th International Congress of Phonetic Sciences*, vol. 1, pp. 154-157. Stockholm.
Horne, M., Strangert, E. & Heldner, M. (1995) Prosodic boundary strength in Swedish: final lengthening and silent interval duration. In *Proceedings of 13th International Congress of Phonetic Sciences*, vol. 1, pp. 170-173. Stockholm.
Huggins, A. W. F. (1978) Speech timing and intelligibility. In *Attention and Performance*, VII. (J. Requin, editor). Hillsdale, N.J.: Erlbaum.
Hunnicut, S. (1985) Intelligibility versus redundancy - Conditions of dependency, *Language and Speech,* **28**, 45-56.

Ingram, D. (1989) *First language acquisition: Method, description and explanation.* Cambridge, UK: Cambridge University Press.

Johns-Lewis, C. (editor) (1986) *Intonation in Discourse.* College-Hill Press, Inc.

Kahn, D. (1976) *Syllable-based generalizations in English phonology.* PhD thesis, University of Massachusetts.

Kaiki, N., Takeda, K. & Sagisaka, Y. (1990) Statistical analysis for segmental duration rules in Japanese speech synthesis. In *Proceedings of International Conference on Spoken Language Processing*. Kobe, Japan.

Kent, R. D. & Burkard, R. (1981) Changes in acoustic correlates of speech production. In *Aging: Communication processes and disorder* (D. S. Beasley & G. A. Davis, editors). New York: Grune and Stratton.

Kent, R. D., Adams, S. G. & Turner, G. S. (1996) Models of Speech Production. In *Principles of experimental phonetics* (N. J. Lass, editor), pp. 30-45. USA: Mosby.

Klatt, D. (1975) Vowel lengthening is syntactically determined in a connected discourse, *Journal of Phonetics,* **3**, 332-335.

Klatt, D. (1976) Linguistic uses of segmental duration in English: Acoustic and perceptual evidence, *Journal of the Acoustical Society of America,* **59**(5), 1208-1221.

Klatt, D. & Cooper, W. E. (1975) Perception of segment duration in sentence context. In *Structure and process in speech perception* (A. Cohen & S. Nooteboom, editors), pp. 69-89. Heidelberg: Springer-Verlag.

Knowles, G. (1991) Prosodic labelling: the problem of tone group boundaries, In *English computer corpora: selected papers and research guide* (S. Johansson & A-B. Stenström, editors). Berlin: Mouton de Gruyter.

Kohler, K. (1983) Prosodic boundary signals in German. *Phonetica,* **40**, 89-134.

Kohler, K. (1991) Isochrony, units of rhythmic organization and speech rate. In *Proceedings of 12th International Congress of Phonetic Sciences*, vol. 1, pp. 257-261. Aix-en-Provence.

Kowal, S., O'Connell, D. C. & Sabin, E .F. (1975) Development of temporal patterning and vocal hesitations in spontaneous narratives. *Journal of Psycholinguistic research*, **4**, 195-207.

Kowal, S., Wiese, R. & O'Connell, D. C. (1983) The use of time in storytelling, *Language and Speech*, **26**, 377-392.

Kozhevnikov, V. & Chistovich, L. (1965) Rech' Artikuliacia i vosprilatie. Moscow-Leningrad. (Trans. in *Speech: Articulation and Perception.* Washington: Joint Publications Research Service, 1966, **30**, 543ff.).

Ladd. D. R. (1984) Declination: a review and some hypotheses, *Phonology Yearbook,* **1**, 53-74.

Ladd, D. R. (1986) Intonational phrasing: The case for recursive prosodic structure, *Phonology Yearbook,* **3**, 311-340.

Ladd, D. R. (1992) An introduction to intonational phonology. In *Laboratory Phonology II* (G. J. Docherty & D. R. Ladd, editors), pp. 331-334. Cambridge, UK: Cambridge University Press.

Ladd, R. D. (1996) *Intonational Phonology.* Cambridge, UK: Cambridge University Press.

Ladd, D. R. & Campbell, W. N. (1991) Theories of prosodic structure: Evidence from syllable duration. In *Proceedings of 12th International Congress of Phonetic Sciences*, vol. 2, pp. 290-293. Aix-en-Provence.
Lehiste, I. (1970) *Suprasegmentals*. Cambridge, MA: MIT Press.
Lehiste, I. (1972) Timing of utterances and linguistic boundaries, *Journal of the Acoustical Society of America,* **51**, 2018-2024.
Lehiste, I. (1973) Rhythmic units and syntactic units in production and perception. *Journal of the Acoustical Society of America,* **54**, 1228-1234.
Lehiste, I. (1975) The role of temporal factors in the establishment of linguistic units and boundaries. In *Phonologica 1972* (W. U. Dressler & F. V. Mares, editors). München-Salzburg: Wilhelm Fink Verlag.
Lehiste, I. (1990) Phonetic investigating of metrical structure in orally produced poetry, *Journal of Phonetics*, **18**, 123-133.
Lehtonen, J. (1979) Speech rate and pauses in the English of Finns, Swedish-speaking Finns and Swedes. In *Perception and production of English: papers on interlanguage* (R. Palmberg, editor). Åbo: Publications of the department of English, Åbo Akademi.
Lenneberg, E. H. (1967) *Biological Foundations of Language*. New York: John Wiley.
Levelt, W. J. M. (1989) *Speaking: From intention to articulation.* Cambridge, MA: MIT Press.
Liberman, A. M. (1970) The grammars of speech and language, *Cognitive Psychology,* **1**, 301-323.
Liberman, M.Y. (1975) *The intonational system of English.* PhD thesis. MIT. Reproduced by Indiana University Linguistics Club, 1978, Bloomington, Indiana.
Liberman, M. & Pierrehumbert, J. (1984) Intonational invariance under changes in pitch range and length. In *Language Sound Structure* (M. Aronoff & R. T. Oehrle, editors). MIT.
Lieberman, P. (1967) *Intonation, perception, and language.* Cambridge: MIT Press.
Lindblom, B. (1963) Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America*, **35**, 1773-1781.
Lindblom, B. & Rapp, K. (1973) Some temporal regularities of spoken Swedish, *Papers from the Institute of Linguistics,* **21**, pp.1-59. University of Stockholm.
Lindblom, B., Lyberg, B & Holmgren K. (1981) Durational patterns of Swedish phonology: Do they reflect short memory processes? Bloomington, IN: IULC.
Linton, M. & Gallo, P. S. (1975) *The practical statistician: Simplified handbook of statistics*. Monterey, CA: Brooks/Cole Publishing company.
Ludvíková, M. (1987) O slabice. In *O češtině v číslech* (V. Těšitelová, editor), pp. 104-105. Praha.
Lyberg, B. (1981) Temporal properties of spoken Swedish. MILUS, vol. 6, Stockholm University.

Lyberg, B. & Ekholm B (1994) The final lengthening phenomenon in Swedish - a consequence of default sentence accent? In *Proceedings of the 1994 International Conference on Spoken Language Processing (ICSLP),* pp. 135-138.Yokohama.

Maeda, S. (1976). *A characterization of American English intonation.* PhD thesis, MIT.

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

Mason, R. D., Lind, D. A. & Marchal, W. G. (1983) *Statistics: An introduction.* Toledo: Harcourt Brace Jovanovich.

McAllister, J., Potts, A. & Mason, K. (1994) Word duration in monologue and dialogue speech, *Language and Speech,* **37**, 393-405.

Miller, J. L. (1981) Effects of speaking rate on segmental distinctions. In *Perspectives on the study of speech* (P. D. Eimas & J. L. Miller, editors), pp. 39-74. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Miller, J., Grosjean, F. & Lomanto, C. (1984) Articulation rate and its variability in spontaneous speech: A reanalysis and some implications, *Phonetica*, **41**, 215-225.

Nakatani, L. H., O'Connor, K. D. & Aston, C. H. (1981) Prosodic aspects of Americal English speech rhythm, *Phonetica,* **38**, 84-106.

Nooteboom, S.G. (1991) Some observations on the temporal organization and rhythm of speech. In *Proceedings of 12th International Congress of Phonetic Sciences,* vol. I, pp. 228-237. Aix-en-Provence.

Nooteboom, S. G. & Eefting, W. (1994) Evidence for the adaptive nature of speech on the phrase level and below, *Phonetica,* **51**, 92-98.

O'Connel, D. C. & Kowal, S. (1983) Pausology. In *Computers in Language Research 2* (W. A. Sedelow, Jr. & S. Yeates Sedelow, editors), pp. 221-301. Berlin-New York: Walter de Gruyter & Co.

O'Connor, J. D. & Arnold, G. F. (1973) *Intonation of colloquial English.* London: Longman.

Ofuka, E. (1996) *Acoustic and perceptual analyses of politeness in Japanese speech.* Unpublished PhD thesis. University of Leeds.

Ohala, J. & Ewan, W. G. (1973) Speed of pitch change. *Journal of the Acoustical Society of America,* **53**, 345ff.

Olive, J. (1974) Speech synthesis by rule. *Speech Communication Seminar,* Stockholm, August 1-3.

Oller, D. K. (1973) The duration of speech segments: the effect of position in utterance and word length. *Journal of the AcousticalSociety of America,* **54**, 1235-1273.

Ondráčková, J. (1954a) O mluvním rytmu v češtině I, *Slovo a slovesnost,* **15**, 24-29.

Ondráčková, J. (1954b) O mluvním rytmu v češtině II, *Slovo a slovesnost,* **15**, 145-157.

O'Shaughnessy, D. (1976) *Modelling fundamental frequency, and its relationship to syntax, semantics, and phonetics.* Unpublished PhD thesis. MIT.

O'Shaughnessy, D. (1995) Timing patterns in fluent and disfluent spontaneous speech. In *Proceedings of the International Conference on Acoustics, Speech and Signal processing,* pp. 600-603. Detroit, Michigan, USA.

Osser, H. & Peng, F. (1964) A cross cultural study of speech rate. *Language and Speech*, **7**, 120-125.

Paccia, J. M. & Cooper, W. E. (1981) The processing of phrase structures in speech production. In *Perspectives on the study of speech* (P. D. Eimas & J. L. Miller, editors), pp. 311-336. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Palková, Z. (1974) *Rytmická v"stavba prozaického textu.* Praha: Academia.

Palková, Z. (1984) Syntactic dispositions as a factor in discourse segmentation. In *Proceedings of the 10th International Congress of Phonetic Sciences.* Utrecht.

Palková, Z. (1987) Einige Beziehungen zu linearen Eigenschaften der Phonischen Einheiten im Spontangespräch. In *Probleme und sprachwissenschaftliche Methoden der Kommunikationsbefähigung* (E. Stock, editor), pp. 75-80. Wiss. Beiträge, 51, Univ. Halle/Saale.

Palková, Z. (1994) *Fonetika a fonologie češtiny.* Praha: Karolinum.

Palková, Z. & Ptáček, M. (1995) Rhythmical segmentation of synthetic speech. Paper at the Colloquium on Theories of Circuits, Systems and Signals. Charles University, Prague.

Pavel, O. (1974) *Jak jsem potkal ryby.* Prague: Mladá fronta.

Petřík, S. (1938) O hudební stránce středočeské věty. Praha.

Petr, J. (editor) (1986) *Mluvnice češtiny I-III.* Praha: Academia.

Pierrehumbert, J. (1980) *The phonology and phonetics of English intonation.* PhD thesis, MIT. Published by Indiana University Linguistics Club.

Pierrehumbert, J. (1981) Synthesizing intonation, *Journal of the Acoustical Society of America*, **70**(4), 985-995.

Pitrelli, J. F., Beckman, M. E. & Hirschberg, J. (1994) Evaluation of prosodic transcription labelling reliability in the ToBI framework. In *Proceedings of the 1994 International Conference on Spoken Language Processing (ICSLP),* pp. 123-126. Yokohama.

Quirk, R. & Greenbaum, S. (1973) *A university grammar of English.* London: Longman.

Radford, A. (1988) *Transformational grammar: A first course*: Cambridge, UK: Cambridge University Press.

Ratkowsky, D. A. (1989) *Handbook of non-linear regression models.* Washington: Pullman.

Raupach, M. (1980) Temporal variables in first and second language speech production. In *Temporal variables in speech: Studies in honour of F. Goldman-Eisler* (H. W. Dechert & M. Raupach, editors). The Hague: Mouton Publishers.

Rees, M. (1975) The domain of isochrony. *Work in Progress*, **8**, pp. 14-28. Edinburgh University, Department of Linguistics.

Rietveld, A. C. M. & Eefting, W. (1988) Temporal variation in natural speech. In *Proceedings*, **12**, pp. 12-13. Department of Language and Speech, Phonetics section. University of Nijmegen.

Roach, P. (1982) On the distinction between 'stress-timed' and 'syllable-timed' languages. In *Linguistic controversies* (D. Crystal, editor), pp. 73-79. London: Edward Arnold.

Roach, P. (1994) Conversion between prosodic transcription systems: 'Standard British' and ToBI. *Speech Communication*, **15**, 91-99.
Roach, P. (in press) Some languages are spoken more quickly than others. In *Language myths* (P. Trudgill & L. Bauer, editors).
Roach, P., Arnfield, S. & Hallum, E. (1996) BABEL: A multi-language database. In *Proceedings of the Australian International Conference on Speech Science and Technology (SST 96)*, pp. 351-354.
Rochester, S. R. (1973) The significance of pauses in spontaneous speech, *Journal of Psycholinguistic Research*, 2, 51-81.
Romportl, M. (1958) *Zvuková stránka souvislé řeči v nářečích na Těšínsku.* Ostrava.
Romportl, M. (1973) *Studies in Phonetics.* Praha.
Romportl, M. (1973) Základy fonetiky. Praha.
Sabin, E. J., Clemmer, E, J., O'Connell, D. C. & Kowal, S. (1979) A pausological approach to speech development. In *Of Speech and Time: Temporal Speech Patterns in Interpersonal Contexts* (A.W.Siegman & S. Feldstein, editors), pp. 35-55. Hillsdale, NJ: Lawrence Erlbaum.
Sanderman, A. A (1995) Phonetic realisation of prosodic boundaries in synthetic speech. In *Proceedings of 13th International Congress of Phonetic Sciences*, vol. 2, pp. 342-345. Stockholm.
Sanderman, A. A. & Collier, R. (1994) Prosodic phrasing at the sentence level. *Festschrift for K. Harris of Physics, Modern Acoustics and Signal Processing Series.* American Institute of Physics.
Schuetze-Coburn, S., Shapley, M. & Weber, E. G. (1991) Units of intonation in discourse: A comparison of acoustic and auditory analyses, *Language and Speech,* **34**, 207-234.
Sedláková, J. (1989) *Variability of speech tempo in monologues.* Unpublished undergraduate thesis. Charles University, Prague.
Selkirk, E. (1980a) The role of prosodic categories in English word stress, *Linguistic Inquiry,* **11**, 563-605.
Selkirk, E. (1980b) Prosodic domains in phonology: Sanskrit revisited. In *Juncture,* Saratoga (M. Aronoff & M. L. Kean, editors). Ca.: Anma Libri.
Selkirk, E. (1984) *Phonology and syntax: the relation between sound and structure.* Cambridge, MA: MIT Press.
Selkirk, E. (1995) Sentence prosody: Intonation, stress, and phrasing. In *The handbook of phonological Theory* (J. A. Goldsmith, editor), pp. 550-569. Oxford: Blackwell.
Shields, L. & Balota, D. (1991). Repetitions and associative context effects in speech production, *Language and Speech,* **31**, 135-156.
Silverman, K., Beckman, M., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., Pierrehumbert, J. & Hirschberg, J. (1992) ToBI: A standard for labelling English prosody. In *Proceedings of the 1992 International Conference on Spoken Language Processing (ICSLP)*, vol. 2, 867-870. Banff, Alberta, Canada.
Skaličková, A. (1954) K otázce podstaty slabiky. In *Slovo a slovesnost,* **15**, 19-24.

Slis, I. H., Haselager, G. J. T. & Rietveld, A. C. M. (1988) Development of speech rate in Dutch children. *Proceedings*, **12**, Department of Language and Speech, Phonetics section, University of Nijmegen.

Smith, B. C., Wasowicz, T. & Preston, T. (1987) Temporal characteristics in the speech of normal elderly adults, *Journal of Speech and Hearing Research,* **30**, 522-528.

Sorensen, J. M. & Cooper, W. E. (1980) Syntactic coding of fundamental frequency in speech production. In *Perception and production of fluent speech* (R. A. Cole, editor). Hillsdale, N. J.: New Jersey: Lawrence Erlbaum Associates.

Starkweather, C. W. (1980) Speech fluency and its development in normal children. In *Speech and Language: Advances in Basic Research and Practice* (N. J. Lass, editor). New York: Academic Press.

Stenström (1990) Pauses in Monologue and Dialogue. In *The London-Lund corpus of spoken English: Description and research* (J. Svartvik, editor), Lund Univesity Press, Lund.

Strangert, E. (1993) Speaking style and pausing. In *Reports form the Department of Phonetics - PHONUM,* **2**, pp. 121-137. University of Umeå.

Swerts, M. & Geluykens, R. C. (1994) Prosody as a marker of intonation flow in spoken discourse, *Language and Speech,***37**, 21-43.

t'Hart, J., Collier, R. & Cohen, A. (1990) *A perceptual study of intonation.* Cambridge, UK: Cambridge University Press.

Thorsen, N. (1980) A study of the perception of sentence intonation: evidence from Danish, *Journal of the Acoustical Society of America,* **67**, 1014-1030.

Trim, J. L. M. (1959) Major and minor tone-groups in English, *Le Maître Phonétique,* **112**, 26-29.

Uldall, E. (1971) Isochronous stresses in R. P. In *Form and substance* (E. Hammerich, editor). Copenhagen.

Vaane, E. (1982) Subjective estimation of speech rate, *Phonetica,* **39**, 136-149.

Vaissière, J. (1971) Contribution à la synthèse par regles du Français. PhD thesis, Université des Langues et Lettres de Grenoble.

Vaissière, J. (1983) Language-independent prosodic features. In *Prosody: Models and Measurements* (A. Cutler & D. R. Ladd, editors), pp. 53-66. Heidelberg: Springer-Verlag.

Vaissière, J. (1991) Perceiving rhythm in French? In *Proceedings of 12th International Congress of Phonetic Sciences*, vol. 4, pp. 258-261. Aix-en-Provence.

van Santen, J. P. H. (1992) Contextual effects on vowel duration, *Speech Communication,* **11**(6), 513-546.

Vincent, M., Di Christo, A. & Hirst, D. (1995) Prosodic features of finality for intonation units in French. In *Proceedings of 13th International Congress of Phonetic Sciences*, vol. 2, pp. 718-721. Stockholm.

Wang, M. Q. & Hirschberg, J. (1992) Automatic classification of intonational phrase boundaries, *Computer Speech and Language,* **6**, 175-196.

Warren, P. (1985) *The temporal organization and perception of speech.* Unpublished PhD thesis, Cambridge University.

Webb, J. T. (1969) Subject speech rates as function of interviewer's behaviour, *Language and Speech*, **12**, 54-67.

Weismer, G. & Ingrisano, D. (1979) Phrase-level timing patterns in English: effects of emphatic stress location and speaking rate, *Journal of Speech and Hearing Research*, **22**, 516-533.

Welford, A.T. (1977) Causes of slowing performance with age. *Interdisciplinary Topics in Gerontology,* **11**, 43-51.

Wenk, B. J. & Wioland, F. (1982) Is French really syllable-timed? *Journal of Phonetics,* **10**, 193-216.

Whalen, D. H. (1990) Coarticulation is largely planned. *Journal of Phonetics,* **18**, 3-35.

Wichmann, A. & Knowles, G. (1995) How determinable are intonation units? In *Proceedings of 13th International Congress of Phonetic Sciences,* vol. 2, pp. 222-225. Stockholm.

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M. & Price, P. (1992) Segmental durations in the vicinity of prosodic phrase boundaries, *Journal of the Acoustical Society of America,* **91**, 1707-1717.

Williams, B. & Hiller, S. M. (1994) The question of randomness in English foot timing: a control experiment, *Journal of Phonetics*, **22**, 423-440.

Zellner, B. (1994) Pauses and the temporal structure of speech. In *Fundamentals of speech synthesis and speech recognition* (E. Keller, editor), pp. 41-62. John Wiley & Sons Ltd.

Zima, P. (1959) O otázce klasifikace mluvního tempa. In *Slovo a slovesnost,* **20**, 96ff.

Oxford U...
Phonetics L...