Zharkova, N. 2007. *An Investigation of Coarticulation Resistance in Speech Production Using Ultrasound.* PhD thesis. Queen Margaret University.

AN INVESTIGATION OF COARTICULATION RESISTANCE IN SPEECH PRODUCTION USING ULTRASOUND

NATALIA ZHARKOVA

A thesis submitted in partial fulfilment of the
requirements for the degree of
Doctor of Philosophy

QUEEN MARGARET UNIVERSITY

2007

“Language is certainly a system,
but it is a system complicated by many controversies”
(L. V. Scherba, 1943)

ABSTRACT

Sound segments show considerable influence from neighbouring segments, which is described as being the result of coarticulation. None of the previous reports on coarticulation in vowel-consonant-vowel (VCV) sequences has used ultrasound. One advantage of ultrasound is that it provides information about the shape of most of the midsagittal tongue contour. In this work, ultrasound is employed for studying symmetrical VCV sequences, like /ipi/ and /ubu/, and methods for analysing coarticulation are refined. The use of electropalatography (EPG) in combination with ultrasound is piloted in the study. A unified approach is achieved to describing lingual behaviour during the interaction of different speech sounds, by using the concept of Coarticulation Resistance, which implies that different sounds resist coarticulatory influence to different degrees.

The following research questions were investigated: how does the tongue shape change from one segment to the next in symmetrical VCV sequences? Do the vowels influence the consonant? Does the consonant influence the vowels? Is the vocalic influence on the consonant greater than the consonantal influence on the vowels? What are the differences between lingual and non-lingual consonants with respect to lingual coarticulation? Does the syllable/word boundary affect the coarticulatory pattern? Ultrasound data were collected using the QMUC ultrasound system, and in the final experiment some EPG data were also collected. The data were Russian nonsense VCVs with /i/, /u/, /a/ and bilabial stops; English nonsense VhV sequences with /i/, /u/, /a/; English /aka/, /ata/ and /iti/ sequences, forming part of real speech.

The results show a significant vowel influence on all intervocalic consonants. Lingual consonants significantly influence their neighbouring vowels. The vocalic influence on the consonants is significantly greater than the consonantal influence on the vowels. Non-lingual consonants exhibit varying coarticulatory patterns. Syllable and word boundary influence on VCV coarticulation is demonstrated. The results are interpreted and discussed in terms of the Coarticulation Resistance theory: Coarticulation Resistance of speech segments varies, depending on segment type, syllable boundary, and language. A method of quantifying Coarticulation Resistance based on ultrasound data is suggested.

ACKNOWLEDGEMENTS

This work would have never appeared without my director of studies Nigel Hewlett. For these three years, he has been not only a senior colleague and a supervisor for me, but an experienced and helpful guide in the scientific world and in real life.

I am very grateful to James Scobbie, my second supervisor, for being always ready to answer various questions, for a lot of comments and suggestions, for many academic and non-academic discussions, and for his unique spontaneous way of supervising.

I am greatly indebted to my examiners, John Local and Fiona Gibbon. They took so much interest in my work, and they were so not-indifferent! They helped me to look at my work from a new angle and to learn about many challenges for the future.

The idea of applying the nearest neighbour algorithm to comparing tongue configurations belongs to Nigel Hewlett. I am grateful to Nigel Hewlett and to Yolanda Vazquez Alvarez for development of the scripts I have used and further adapted for my work.

I would like to thank two people for technical support. Steve Cowen has always been there for any technical help, or for my crying on his shoulder when something beyond control would happen when least needed. Yolanda Vazquez Alvarez assisted me in recording experimental subjects and helped with Matlab and with statistical analysis.

I express special thanks for help with statistics to Ivan Yuen, Alice Lee, Robert Rush (Centre for Integrated Healthcare Research, QMUC), Robin Lickley, Don Sutherland, and Maria Wolters.

A great thank you to all my subjects, who were ever so patient, kind and friendly during the non-trivial ultrasound (and EPG for some!) recording procedures.

Many thanks to Bill Hardcastle, for helping me out with finding relevant literature several times when I had questions I thought I would never find answers to.

Thanks a lot to Alan Wrench, for advice on the technical side of data analysis, and for answering countless linguistic and extra-linguistic questions.

I would like to thank Olga Gordeeva for a lot of practical help and advice, from the induction throughout the PhD work.

I am deeply grateful to all the colleagues in Speech and Hearing Sciences, for all possible ways of support, answers, encouragement, sharing tea and coffee, and for that special atmosphere of friendliness I so appreciate.

And I would like to thank my mother so much, for everything!

CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

Coarticulation

Influence that sound segments have on other segments, and experience from other segments, while interacting in the process of speech.

Coarticulation Resistance (CR)

Ability of speech sounds to maintain the same articulatory/acoustic properties across different contexts.

Coarticulation Resistance Coefficient (CRC)

A coefficient introduced in this work, representing Coarticulation Resistance of a speech sound in relation to a particular context, calculated from ultrasound data.

Degree of Articulatory Constraint (DAC)

A schematic representation of Coarticulation Resistance of a speech sound, based on the degree of constraint on the tongue dorsum for the production of that sound.

Gesture

A coordinated articulatory movement that is dynamically specified, i.e., it has its own intrinsic temporal dimension.

Inter-utterance Speech rest Position (ISP)

General posture of the articulatory organs typical of a language, or a neutral position of the articulatory organs.

Nearest Neighbour technique

A technique designed for comparison of different tongue curves. It involves comparing each point on one curve to each point on the other curve and finding the shortest distance, called the nearest neighbour distance, then averaging nearest neighbour distances, and arriving at one nearest neighbour distance to represent the average distance between the curves.

Spline

A smooth curve generated with a mathematical formula, which passes through a number of pre-specified points.

Trough

A term meaning discontinuity in coarticulation between the two vowels in a VCV sequence with symmetrical vowels (e.g., tongue deactivation during a bilabial consonant between the two vowels, or diminution of lip protrusion in a lingual consonant between two rounded vowels).

MAIN ABBREVIATIONS USED IN THIS TEXT

QMUC	Queen Margaret University College
CR	Coarticulation Resistance
DAC	Degree of Articulatory Constraint
CRC	Coarticulation Resistance Coefficient
S1	Subject 1
S2	Subject 2
S3	Subject 3
V	vowel
C	consonant
VCV	vowel-consonant-vowel sequence, with phonologically identical vowels
VhV	vowel-/h/-vowel sequence, with phonologically identical vowels
V1	the first vowel of a VCV sequence
V2	the second vowel of a VCV sequence
V1 spline	the spline, in the ultrasound analysis software, made at the V1 annotation point
V1 curve	the curve, specified by a set of points exported to Matlab, corresponding to the V1 spline
V2 spline	the spline, in the ultrasound analysis software, made at the V2 annotation point
V2 curve	the curve, specified by a set of points exported to Matlab, corresponding to the V2 spline
C spline	the spline, in the ultrasound analysis software, made at the C annotation point
C curve	the curve, specified by a set of points exported to Matlab, corresponding to the C spline
a1	V1 of an /aCa/ sequence

Abbreviations

a2	V2 of an /aCa/ sequence
i1	V1 of an /iCi/ sequence
i2	V2 of an /iCi/ sequence
k_a	/k/ in /aka/
t_a	/t/ in /ata/
t_i	/t/ in /iti/
a_k	/a/ in /aka/
a_t	/a/ in /ata/
$a1_k$	/a1/ in /aka/
$a2_k$	/a2/ in /aka/
$a1_t$	/a1/ in /ata/
$a2_t$	/a2/ in /ata/
$i1_t$	/i1/ in /iti/
$i2_t$	/i2/ in /iti/

1. INTRODUCTION

1.1. Rationale for this study

It is a well-known linguistic fact that "sound segments are highly sensitive to context and show considerable influence from neighbouring segments. Such contextual effects are described as being the result of overlapping articulation or coarticulation" (Hardcastle & Hewlett 1999, p. 1). The term "coarticulation" in linguistic studies commonly implies overlapping articulation, or an "articulation which takes place involving in a simultaneous or overlapping way more than one point in the vocal tract" (Crystal 1997, p. 66). This overlapping articulation involves mutual influence of neighbouring sounds in speech, cf. the definition from a popular textbook in linguistics: "coarticulation – the transfer of phonetic features to adjoining segments to make them more alike, e.g., vowels become [+ nasal] when followed by consonants that are [+ nasal]" (Fromkin et al. 2003, p. 577). So many acoustic, articulatory and perceptual characteristics of speech sounds can be considered to arise from coarticulatory processes happening in speech, that it is not surprising to see the following observation in a recent experimental study: "There are probably few phonetic topics that have generated so much experimentation and theoretical speculation as coarticulation. Nonetheless, it is fair to say that, despite massive research efforts, our understanding of the physiological origins and perceptual motivation for this phenomenon still remains rather incomplete" (Ericsdotter et al. 1999, p. 1885).

One of the most interesting issues about coarticulation for linguists is the interplay of invariant units and their variable realisations: "it is essential to the concept of coarticulation that at some level there be invariant, discrete units underlying the variable and continuous activity of speech production" (Kühnert & Nolan 1999, p. 7). The current study assumes the existence of invariant units and tries to find out more about their identity and about the variability of their realisations in speech, by studying articulatory interaction of neighbouring speech sounds.

Different types of contextual dependencies have been reported in the literature. For example, vowels have been shown to vary more depending on their position in relation to stress and prosodic boundaries than on the neighbouring consonants, while most of the consonantal variation is due to the vocalic context, though position in word/phrase matters too, of course (e.g., Keating et al. 1994). An example of consonantal variation dependent on the vowel context is tongue behaviour during intervocalic non-lingual consonants. The tongue position during labial consonants or /h/ has been shown to vary greatly, adopting the tongue position of neighbouring vowels (e.g., Recasens 1999), as for the production of these consonants, the tongue does not need to take any particular posture. Lingual consonants' tongue shape differs more from neighbouring vowels, the degree of this difference depending on the place of lingual constriction during the consonant formation (e.g., Recasens et al. 1997). Another factor that plays in the coarticulatory interaction of consonants and vowels is how different or similar the target tongue position is for the particular sounds involved (e.g., Recasens 1999; Fowler & Brancazio 2000). For example, in English, during the /ti/ sequence, the tongue does not change its position as much as during the /ri/ sequence, because of the more conflicting requirements on the tongue position for producing /r/ and /i/.

All the segmental interactions described here are further influenced by suprasegmental characteristics of speech, such as syllable structure, word and phrase boundaries, lexical stress and phrasal accent, and position in the utterance. As Modarresi et al. put it: "It is not readily apparent how one disentangles this web of interactions to arrive at a theoretically coherent account of this overlapping articulatory montage" (Modarresi et al. 2004, p. 292).

1.2. Conceptual framework

One of the notions used in the literature for describing the relation between invariant units and variable realisations in speech production is "coarticulation resistance" (CR). The term was introduced in Bladon and Al-Bamerni (1976). That study will be described in detail in Section 2.1.1. The idea of speech sounds being resistant to coarticulation implies that the sounds retain their phonetic properties across different contexts. This approach to studying coarticulation is based on describing how similar the sounds are under different segmental and suprasegmental influence. The appeal of this approach, to my mind, is that the notion of CR appears to be a way of directly capturing the degree to which variable sounds of speech retain articulatory and acoustic properties of underlying invariant units.

Bladon and Al-Bamerni did not practically implement their ideas for quantifying coarticulatory properties of speech sounds. In later studies, the concept of CR has been used and developed. Currently, the Degree of Articulatory Constraint (DAC; e.g., Recasens et al. 1997; Recasens 2004) model is a well-developed model within the CR approach to coarticulation (see Section 2.1.2 for the detailed description).

The present work is based on the CR approach to coarticulation in speech. The DAC model will be used as a theoretical framework for this research. Throughout this work, the results of the experiments will be interpreted within the DAC model, because this model has a well-developed terminological system for describing coarticulatory processes in speech. However, all the results, discussed within the DAC framework, will be viewed within the larger context of the CR approach, for reasons discussed below.

In the original formulation of the principles of CR by Bladon and Al-Bamerni, there were some very general theoretical predictions, aiming at a theory of speech production. Language-specificity and subject-specificity were mentioned, for example, as possible factors accounting for variation of CR (for a detailed description, see Section 2.1.1). In the DAC model, the idea of CR is the same as the one formulated by Bladon and Al-Bamerni, but the model only aims at describing coarticulation. The proponents of the model call it a “descriptive framework” (Recasens et al. 1997, p. 545). In one of the key publications about the model, it is called “a model of lingual coarticulation” (Recasens et al. 1997, p. 544). The DAC model is narrower than a theory of speech production, in that it aims to describe and predict properties of sounds concerning lingual coarticulation, based on the constraint on tongue dorsum. The DAC model is regarded in this work as a practical application of the CR approach to speech production, designed for studying and describing articulatory characteristics of speech sounds in relation to neighbouring sounds.

This limitation of the DAC model has a positive side: it makes the model very convenient for use with ultrasound data on midsagittal tongue configurations, because the model focuses on lingual coarticulatory properties. Ultrasound data have never been used to test the applicability of the model. So we have grounds for testing the model's theoretical predictions with the new type of data, for obtaining new information on articulatory characteristics of speech sounds, and eventually for a more accurate description of the mechanisms responsible for the interaction of sounds during articulation.

In this study, the DAC model terminology and conceptual framework will be used to test the claims that were formulated in a very general way within the original CR approach, and also the claims that were later formulated more precisely and tested with experimental phonetic methods within the DAC model. The study will analyse different segments, contextual influence and syllable boundary influence, inter-speaker and cross-linguistic variation. Some changes in the framework of the DAC model will be introduced, for describing the results obtained in the experiments. Numerical values characterising resistance of speech sounds to coarticulation will be calculated, based on the results of the ultrasound experiments. The changes introduced will be based on both CR notions and the DAC model terms, and will aim to offer a more accurate description of articulatory characteristics of speech sounds than has hitherto been done.

The terms "Coarticulation Resistance" and "Degree of Articulatory Constraint" are used as defined in the Glossary. The term "degree of CR" is used occasionally, with a meaning synonymous to "DAC", when describing and discussing articulatory constraints not described in the publications within the DAC model. The term "Coarticulation Resistance Coefficient" is introduced in Chapter 7, where suggestions for quantifying degrees of CR based on ultrasound data are presented.

In most of this work, especially when presenting and discussing ultrasound results, the term "coarticulation" is synonymous to "lingual coarticulation". Often, the adjective "lingual" is used explicitly, for making the text clearer.

The abbreviation "VCV" is used in this work meaning a symmetrical vowel-consonant-vowel sequence, i.e., with two phonologically identical vowels. When non-symmetrical sequences are mentioned (e.g., when referring to other studies), the combination of words “vowel-consonant-vowel sequence” is used.

1.3. Method

This study aims to contribute to knowledge in the field of coarticulation in speech, using ultrasound as the principal method of investigation. The “web of interactions” (Modarresi et al. 2004) described in Section 1.1 has been investigated by numerous speech researchers, by various methods (acoustic, articulatory and perceptual analysis of speech, and speech modelling). There do not exist numerous studies of CR in speech sounds with ultrasound. The examples referred to in Section 1.1 show that tongue position during articulation matters a lot in the acoustic structure and auditory impression of the sounds. So studying lingual coarticulation can give us much more information than is available at the present time, on the processes happening in speech with sounds neighbouring each other. Ultrasound is an excellent method for looking at tongue articulation, as it is safe, non-invasive, and therefore allows for collecting large amounts of data. It is relatively new in speech research, and techniques of tongue contour analysis are still being developed (for more on this, see Section 3.2). This work will present some methodological innovations in the ultrasound analysis of coarticulation.

Electropalatography (EPG) is used in this work to a limited extent, in order to obtain some supporting evidence from the multi-channel system of speech analysis, combining ultrasound, EPG and acoustics. EPG is used in the research with the purpose of integrating ultrasound analysis of CR with previous work in the framework of the DAC model, most of which has been done using EPG (see Section 2.1.2 for details).

1.4. Structure of this thesis

Chapter 2 is a description of the background to this study. The CR approach to speech coarticulation and the DAC model, used as a theoretical framework for this research, will be described in detail in Section 2.1. Motivation for choosing the research questions will be presented there, and also in Section 2.2, where studies of tongue behaviour in

VCV sequences with non-lingual consonants are described. In Section 2.3, several theories of coarticulation are presented, which are relevant for the present research. These theories are compared with the DAC model, and the reasons for preferring the latter model to the others are given.

Chapter 3 is concerned with methodological issues. Particular attention is paid to ultrasound, the principal technique used in this work. Section 3.2 contains a description of the technical details involved in the procedure of ultrasound scanning of the tongue, and a short background to the use of ultrasound in speech research. The QMUC ultrasound system is described in Section 3.3, including hardware, software, and the combined ultrasound and EPG setup. General methodology used throughout this work is presented in Section 3.4, including descriptions of recording techniques and approaches to data analysis common to all the three experiments in this work, and statistical procedures for tongue curve analysis designed during the work.

The study consists of three separate but related experiments. Chapters 4, 5 and 6 describe these experiments. In all these chapters, research questions for individual experiments are formulated, particular methodological aspects unique to each experiment are described, the results are presented, and a discussion of these results is provided. The choice of languages and speech materials was dictated by the theoretical framework: speech sounds exhibiting different degrees of resistance to lingual coarticulation were studied. In Chapter 4, Experiment 1 is described. This experiment was aimed at studying coarticulatory processes in Russian VCV sequences with bilabial consonants. The Russian data were compared with existing data on lingual coarticulation in VCVs with bilabial consonants in other languages. Experiment 2, presented in Chapter 5, was designed to study British English symmetrical VhV sequences. Experiment 3, described in Chapter 6, was focused on VCV sequences with lingual consonants. In that experiment, some EPG data were analysed, in addition to ultrasound data. In each of these three chapters, the results are discussed in terms of the DAC model, by interpreting the coarticulatory behaviour of the investigated speech sounds using the concept of CR.

In Chapter 7, a general discussion is provided of all three experiments, unifying their results within the CR approach. A critical evaluation of the DAC model is offered, as related to the results of this work. A proposal for quantifying CR of speech sounds based on ultrasound midsagittal data is presented. The methods used in this work are discussed, and their applicability to the analysis of coarticulation is evaluated.

2. BACKGROUND

"Nemo solus satis sapit" (Maccius Plautus)

2.1. Coarticulation Resistance approach to studying coarticulation

2.1.1. The concept of coarticulation resistance (Bladon & Al-Bamerni)

The work by Bladon and Al-Bamerni (1976), mentioned in Section 1.2, aimed at studying coarticulatory properties of the English /l/. When discussing the differences in coarticulatory behaviour of the three main allophones of /l/ in RP, the authors claimed that their results could be best accounted for "by postulating an articulatory control principle of 'coarticulation resistance'" (Bladon & Al-Bamerni 1976, p. 149). To account for different degrees of coarticulation admitted by different allophones in their data, the researchers suggested that each allophone should be assigned a value for CR, "by rules which may in some instances be language-particular and in others quasi-universal" (Bladon & Al-Bamerni 1976, p. 149). The authors suggested that the CR value may be represented as a numerical coefficient, but they did not go as far as assigning coefficients to the sounds they studied. Below, their study is described, in order to set the context for this research, and to outline the questions that are interesting and deserving of exploration.

Bladon and Al-Bamerni write that CR is "a uniform control principle upon whose information the speech encoding mechanism continuously draws" (Bladon & Al-Bamerni 1976, p. 149). This principle "is associated with speech segments in the form of values whose magnitude varies" (Bladon & Al-Bamerni 1976, p. 149), increasing, in the sounds studied by Bladon and Al-Bamerni, from the clear [l] to the dark [ɫ], and being greater in syllabic than in non-syllabic [ɫ]. The researchers suggest that each of the three main allophones of /l/ "is associated with a different numerical specification for the feature CR, coarticulation resistance" (Bladon & Al-Bamerni 1976, p. 149).

The authors make some generalisations about the principle of CR, based on the existing research at that time. For example, they say that at least some rules for assignment of CR values must be context-sensitive. This claim is based on the observation that voiceless plosives generally are less resistant to coarticulation than voiceless fricatives, but that they become more resistant in consonant clusters with the following /l/. When discussing the results of Amerman et al. (1970), Bladon and Al-Bamerni make the following claim about CR of /s/. The fact that the presence of /s/ in the consonant string preceding an /æ/ in English impeded anticipatory coarticulation, means, according to Bladon and Al-Bamerni, that "the coarticulation resistance specification of /s/ must, at least in those contexts, be high" (Bladon & Al-Bamerni 1976, p. 150). Another example given by Bladon and Al-Bamerni comes from Ladefoged (1967), where a difference was noticed between /k/ in French and in English: before /i/, this consonant is articulatorily advanced in both languages; word-finally, in French it is also advanced, but not in English. Bladon and Al-Bamerni then suggest that English "has a context-sensitive assignment of greater CR to final than to initial velars (and perhaps, not only velars)" (Bladon & Al-Bamerni 1976, p. 150).

An interesting issue about CR is that, according to Bladon and Al-Bamerni (1976), it seems to be idiolectal. As supporting evidence to this claim, the authors refer to the study by Su et al. (1974), where degrees of coarticulated vowel quality in nasals were shown to have a speaker-identifying function. On the other hand, Bladon and Al-Bamerni are quite confident in claiming that some specifications of CR may well be universal. One of them is suggested by the authors to be a high CR value attached to an intonation-group boundary. Another one was a "moderately high" CR specification associated with the left boundary of a CV-type syllable.

In the concluding passage of their article, Bladon and Al-Bamerni say that "coarticulation resistance specifications are not necessarily language-universal; and how far they could be so becomes an interesting question for future research" (Bladon & Al-Bamerni 1976, p. 150).

The concept of CR has been used and developed in more recent studies. The Degree of Articulatory Constraint (DAC) model described in Section 2.1.2 is largely based on the notion of CR.

2.1.2. Degree of Articulatory Constraint model

The Degree of Articulatory Constraint (DAC) model is a dynamically oriented model of coarticulation, in which “phonetic segments are characterised in terms of gestures” (Recasens et al. 1997, p. 544). The DAC model “is based on the assumption that articulatory gestures associated with consecutive segments are coproduced and overlap to different degrees depending on their spatiotemporal properties as well as on prosodic factors and speech rate” (Recasens 2002a, p. 2828). According to the DAC model, “coarticulatory patterns in VCV sequences are determined to a large extent by the production demands for the intervocalic consonant” (Recasens et al. 1998, p. 54). The model introduces degrees of articulatory specification, calling them articulatory constraints. According to the model, an articulator is constrained when it is involved in the formation of a closure or constriction. On the other hand, the constraint is much weaker, or there may even be no constraint “in the case of an articulator which does not intervene in the achievement of an articulatory target (e.g., the tongue dorsum during the production of a labial consonant)” (Recasens 1987, p. 299).

Recasens and his colleagues (e.g., Recasens et al. 1997) propose the term “degree of articulatory constraint” (DAC). This term is one of the central concepts in the model. The degree of articulatory constraint depends on the degree of tongue dorsum constraint during production of particular vowels and consonants, and it determines the degree of the segment’s resistance to coarticulation. In Recasens et al. (1997), based on data from Catalan symmetrical and non-symmetrical vowel-consonant-vowel sequences, the authors propose three DAC values, ranking phonetic sound categories from maximally to minimally constrained in the following way: “/ʃ/, /ɲ/, /i/, /k/, dark /l/, (/s/) (DAC = 3) > /n/, /a/, (/s/) (DAC = 2) > /p/, /ə/ (DAC = 1)” (Recasens et al. 1997, p. 545). The consonant /s/ gets into two categories because on the one hand, the tongue dorsum is subject to coupling effects with the tongue blade during this alveolar consonant production, but on the other hand, the manner of articulation of /s/, i.e., the precise formation of a medial groove for fricatives, “should render /s/ more constrained than nonfricative alveolars” (Recasens et al. 1997, p. 545).

By assigning a DAC value to the segment, the model predicts the degree of the tongue dorsum involvement during the production of this sound. Recasens (2002a) claims that the dorsum of the tongue “is the articulator about which the model can make theoretical predictions so far” (Recasens 2002a, p. 2828).

A few examples illustrate the reasoning of the authors in assigning DAC values for the tongue dorsum to particular sounds. Recasens et al. (1997) claim that the vowel /i/ and the consonant /k/ have the highest degree of articulatory constraint value, DAC = 3. The authors explain it by the considerable involvement of the tongue dorsum in their production. They assign an intermediate DAC value (DAC = 2) to the vowel /a/ and to alveolar consonants, and explain it by the fact that “the tongue dorsum is not directly involved in closure or constriction formation but is subject to coupling effects with the primary articulator” (Recasens et al. 1997, p. 545). In the case of alveolar consonants, these coupling effects involve tongue blade raising causing some tongue dorsum raising to occur; note, however, the arguments about /s/ presented above, and also the fact that the production of the dark /l/ involves “active tongue postdorsum retraction for a secondary dorsopharyngeal constriction” (Recasens et al. 1997, p. 545). In the case of /a/, the tongue root retraction gesture for this vowel brings about some concomitant tongue dorsum lowering. Describing the difference in the DAC values between the consonants not involving tongue dorsum as a principal articulator for their production, Recasens (2002b) writes that “coupling between the tongue dorsum and the primary tongue front articulator causes dentals and alveolars to be more constrained than bilabials” (Recasens 2002b, p. 83). In general, according to the DAC model, consonant coarticulatory sensitivity to the influence of neighbouring vowels varies inversely with the strength of the consonantal effects on vowels, and with the consonant’s DAC value.

The DAC model is largely based on EPG and acoustic studies (e.g., Recasens 1984; Recasens 1987; Recasens et al. 1997; Recasens et al. 1998; Recasens & Pallarès

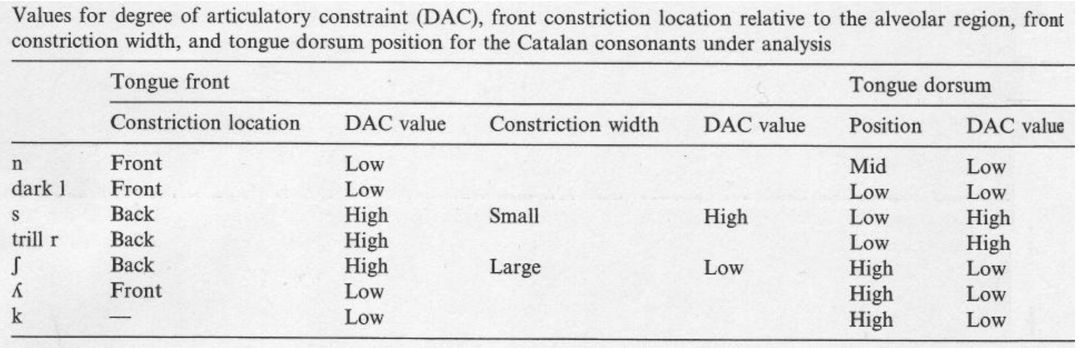
1999; Recasens & Pallarès 2000; Recasens & Pallarès 2001; Recasens 2002c; Recasens 2004). In Recasens (2002a) and Recasens (2002b), some tongue and jaw coarticulation data studied by electromagnetic articulography (EMA) are presented.

The typical way of reporting the data in EPG studies within the DAC model is based on the contact indices data reduction method (e.g., Recasens et al. 1997; Recasens et al. 1998; Recasens & Pallarès 1999; Recasens & Pallarès 2001; Recasens 2004). Lingual contact indices introduced by these researchers represent amount of lingual contact with the palate at different places of articulation. These indices are considered by the researchers to be a method of presenting the data, rather than a way of representing CR of speech sounds. The scale for representing degrees of CR is schematic, and it only has three values, as described above.

Recasens et al. (1997) say that "it needs to be emphasised that this is a preliminary DAC classification which could be improved with a more accurate formulation of the articulatory constraints for consonants and vowels" (Recasens et al. 1997, p. 545). And indeed, in later works, the criteria for defining DAC values for vowels and consonants change. As described in Recasens (2004), degrees of articulatory constraint for phonetic segments are imposed not only on the tongue dorsum, but on another tongue region as well – namely, "tongue front" (or alveolar region). Values for degrees of articulatory constraint are not any more stated in numbers, but rather in two categories: "low" versus "high". Also, constriction location and constriction width are added for describing tongue front DACs, and position is added for tongue dorsum. Tongue dorsum position has three possible values: "high", "mid" or "low" (see Table 2-1).

Another change in the DAC model after Recasens et al. (1997) is the following. Recasens (2004) claims that "the DAC value at the tongue dorsum may vary depending on whether consonants are embedded in clusters or in VCV sequences" (Recasens 2004, p. 437). While DAC values stated in Recasens et al. (1997) are assigned to the sounds from vowel-consonant-vowel sequences, in consonant clusters the DAC values are somewhat different. For example, alveopalatals (/ʃ/, /ʎ/) and velars (/k/) are assigned a high DAC value in vowel-consonant-vowel sequences, and a low DAC value in consonant clusters. Recasens states that "this contrast is in agreement with the finding that these consonants are less sensitive to tongue dorsum effects exerted by vowels than to those exerted by consonants involving active tongue dorsum lowering (i.e., /s, r/)" (Recasens 2004, p. 437).

The table with consonant DAC values in consonant clusters is given in Table 2-1. This table was published in Recasens (2004), and includes several Catalan consonants.



Values for degree of articulatory constraint (DAC), front constriction location relative to the alveolar region, front constriction width, and tongue dorsum position for the Catalan consonants under analysis

	Tongue front				Tongue dorsum	
	Constriction location	DAC value	Constriction width	DAC value	Position	DAC value
n	Front	Low			Mid	Low
dark l	Front	Low			Low	Low
s	Back	High	Small	High	Low	High
trill r	Back	High			Low	High
ʃ	Back	High	Large	Low	High	Low
ʎ	Front	Low			High	Low
k	—	Low			High	Low

Table 2-1. Degree of Articulatory Constraint values for several Catalan consonants from consonant clusters (after Recasens 2004).

In the more recent version of the DAC model (e.g., Recasens 2004), as compared with earlier works, position in the syllable is accepted to be a factor influencing DAC values of consonants. This theoretical innovation in the model is in agreement with the predictions by Bladon and Al-Bamerni (1976) referred to in Section 2.1.1, that CR specifications of speech sounds should depend on syllable structure.

The DAC model has received support among the researchers working on speech motor control aspects of coarticulation. For example, in Modarresi et al. (2004), the authors refer to the DAC model as "currently the only conceptualisation scheme that directly addresses bidirectional coarticulatory effects" (p. 292) of both vowels and consonants. Fowler (2005) is a perceptual study synthesising the original concept of CR and the modern DAC model principles, and applying them to speech perception. Fowler reports the results of the listener's perception of /ə/CV sequences with the second vowel stressed and the consonants with different DAC values. The author gives acoustic and articulatory evidence that the magnitude of the influence of the stressed vowel on anticipatory coarticulation in schwa is modulated by coarticulation resistance, and concludes that the outcome of the perception experiment “is generally consistent with a hypothesis that listeners to speech ‘parse’ the acoustic signal along coarticulatory or phonetic gestural lines and that success in parsing varies with the amount of acoustic evidence talkers provide” (Fowler 2005, p. 199).

In this work, the data will be symmetrical VCV sequences. This type of data allows for studying vowel-on-consonant (V-on-C) influence, consonant-on-vowel (C-on-V) influence and the effect of syllable structure on segmental coarticulation. The DAC model has the tools for describing all these phenomena.

One issue that will be addressed in this work is lingual coarticulation in VCV sequences with non-lingual consonants. The DAC model treats bilabial consonants in VCV sequences as specified for the lowest DAC value, as related to lingual position. Labials have a minimally constrained tongue shape among all consonants, so it is expected by the model that the tongue shape during a bilabial consonant will vary greatly according to the identity of the adjacent vowel. It was reported in, e.g., Fowler and Brancazio (2000), using American English articulatory data, that the labial consonants /b/ and /v/ are less resistant to coarticulation than lingual consonants. However, there exist experimental data in the literature, demonstrating that bilabial consonants do demonstrate some resistance to lingual coarticulation (see Section 2.2 for a detailed description). The works within the DAC model have not been focused on this issue. One of the experiments in this study will address the question of how low the degree of CR is in bilabial consonants, whether it can be considered a zero CR, and what it depends on. In the second experiment, another non-lingual consonant, /h/, is studied. This consonant has not been addressed in the works within the DAC model, and the data on lingual coarticulation in VhV sequences will be an important addition to the information on articulatory constraints in the consonants not involving the tongue as the principal articulator.

Lingual consonants are included in the study, because, according to the DAC model, they have higher DAC values than non-lingual consonants, and hence a noticeable effect on the surrounding vowels from lingual consonants is expected.

Measuring not only V-on-C, but also C-on-V coarticulatory effects will allow for producing a unified description of CR properties of consonants and vowels.

2.2. Lingual coarticulation in VCV sequences with non-lingual consonants

2.2.1. "Trough" patterns: evidence from previous studies

In experimental studies of lingual coarticulation, it has been shown that "the degree of coarticulatory sensitivity at tongue regions which are not directly involved in the formation of a closure or constriction is conditioned by whether they are more or less coupled with the primary lingual articulator" (Recasens 1999, p. 89). Also, it has been experimentally shown that "vowel-dependent coarticulatory effects in tongue body activity are larger for labial stops and for labiodental fricatives than for lingual consonants, since this articulatory structure does not intervene in the production of the former consonantal category" (Recasens 1999, p. 89). These claims agree with the DAC model's predictions described in Section 2.1.2.

However, it has also been experimentally demonstrated that the tongue position does not always stay the same throughout a VCV sequence with symmetrical vowels and a bilabial consonant – a discontinuity in lingual coarticulation can occur. This discontinuity needs some explanation, since a continuous tongue movement would arguably be expected between two identical vowels during the bilabial stop, which is a segment which does not require any particular tongue position different from that of the neighbouring vowels, and could fully accommodate its tongue shape to that of any vowel's. The reasons for this discontinuity in lingual behaviour are a matter of debate in the literature.

The discontinuity, or the "trough", in coarticulation in VCV sequences with C being a bilabial stop consonant has been addressed in a number of works[1]. The

---

[1] Besides, deactivation of labial protrusion for /uCu/ utterances was demonstrated by various researchers (e.g., McAllister 1978; Gay 1979; Engstrand 1981; Perkell 1986), and it was also called a "trough" in the

phenomenon has been described in the following ways: tongue lowering during the medial consonant (Houde 1967), two distinct bursts of genioglossus muscle activity for the two vowels (Gay 1974), cessation of genioglossus activity between the vowels (Bell-Berti & Harris 1974), "relaxation of vowel-related tongue position" as evidenced by tongue movement towards neutral position between the vowels (Engstrand 1988), "temporary reduction of tongue height" (Engstrand et al. 1996), "turning off" or diminution of underlying muscle activation for the vowel (Lindblom et al. 2002), "muscular deactivation during the consonantal part of a VCV-sequence" (Fuchs et al. 2004), "slight lowering of the tongue during the C closure phase" (Vazquez Alvarez et al. 2004).

Works touching on this aspect of VCV coarticulation also include, for example, Engstrand (1989), Boyce (1990), McAllister and Engstrand (1991), McAllister and Engstrand (1992), Farnetani and Recasens (1993), Keating et al. (1994), Svirsky et al. (1997), Hewlett et al. (2004), Modarresi et al. (2004), Yuen et al. (2005), Vazquez Alvarez (2006). This discontinuity in tongue position in VCV sequences with bilabial consonants has been registered using numerous experimental phonetic methods, including electromyography (e.g., Bell-Berti & Harris 1974; Gay 1974; Gay et al. 1974; Fuchs et al. 2004), EPG (e.g., McAllister & Engstrand 1991; McAllister & Engstrand 1992; Engstrand et al. 1996), spectrography (e.g., Lindblom et al. 2002), X-ray (Engstrand 1988; Lindblom et al. 2002), EMA (Fuchs et al. 2004), ultrasound (Vazquez Alvarez et al. 2004; Hewlett et al. 2004).

This coarticulatory, or rather, "counter-coarticulatory", phenomenon has been called "trough effect", after the lowering of the tongue between the two vowels of the VCV. In some works, this term ("trough effect") is used: e.g., Lindblom et al. (2002), Fuchs et al. (2004), Vazquez Alvarez et al. (2004), Hewlett et al. (2004). In other works though, mainly earlier ones, only the word "trough" is used (e.g., Bell-Berti & Harris 1981; Perkell 1986; Engstrand et al. 1996). Other works analysing similar VCV sequences and comparable phenomena do not use this terminology at all (e.g., Svirsky et

---

literature. In this chapter, studies of troughs in lip protrusion will be occasionally mentioned, as supporting evidence to the data on troughs in tongue movement.

al. 1997). So in this work, in order to avoid confusion, the term “trough effect” will not be used. In order to signify tongue lowering between two vowels, the term “trough” will be used consistently across this work[2]. Other terms will be introduced in due course for different coarticulatory patterns.

In Sections 2.2.3 – 2.2.12, the studies that found discontinuity in tongue behaviour in VCV sequences will be described, and possible reasons for the observed patterns, as suggested in those studies, will be presented.

2.2.2. Deactivation

The “counter-coarticulatory” phenomenon referred to in Section 2.2.1 has been rather commonly referred to as “deactivation” of the articulators between the two vowels of the VCV sequence. To talk about deactivation of tongue position between the two vowels of a VCV sequence with a non-lingual consonant, it is necessary to answer the question of what deactivation means, and the question of where the tongue goes when it deactivates.

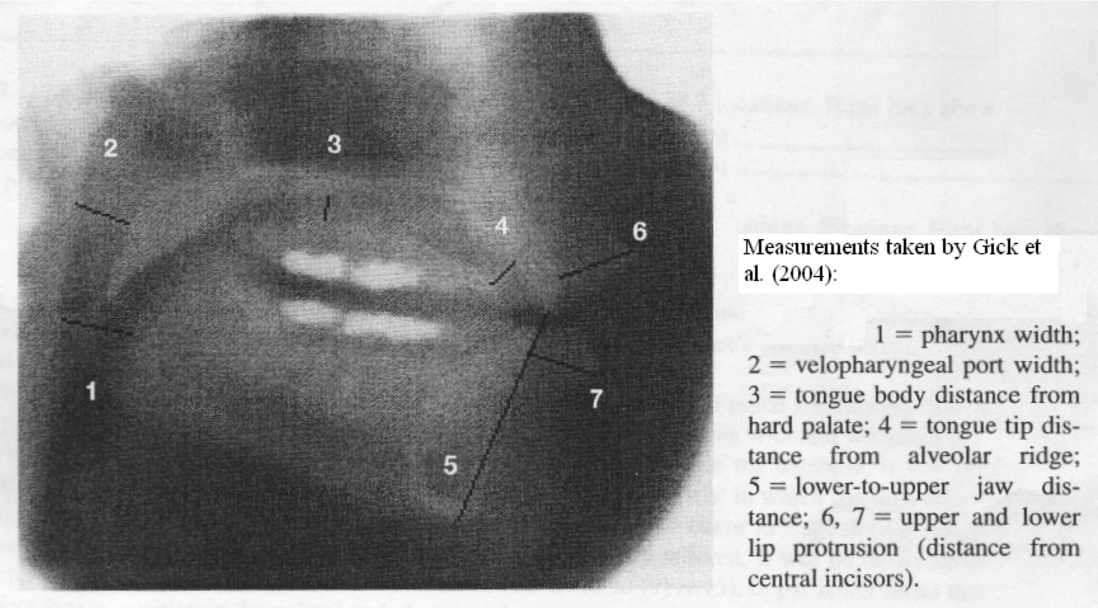
Lindblom et al. (2002) describe the effect occurring in /ibi/ and /ipi/ sequences as “a diminuition or ‘turning off’ of underlying muscle activation for V1, with a concomitant lowering of the tongue dorsum that continues into the stop closure interval” (Lindblom et al. 2002, p. 245). Deactivation thus has to do with muscles, and when it happens, there is an effect on tongue kinematics. It seems logical that when the muscles required for producing particular (especially peripheral) vowels are deactivated, the tongue would move towards the posture which has been called “neutral position” (e.g., Chomsky & Halle 1968; Perkell 1969), or “inter-utterance speech rest position” (Gick et al. 2004)[3]. These terms have been used to describe the general posture of articulatory organs typical of a language, or a neutral position of articulatory organs. Until recently, there have not been many attempts at measuring inter-utterance speech rest position (ISP, as Gick et al., 2004, abbreviate it), its quantitative description and a possible cross-

---

[2] The word “trough” will also be occasionally used in this chapter to signify diminution of lip protrusion in VCV sequences with labialised vowels, in order to be consistent with the use of this term in the literature in this sense (see footnote 1). In these cases, explicit indication will be made that troughs in lip protrusion are meant.

[3] Cf. the term “articulatory settings” (for a historical survey, see, e.g., Wilson 2006).

linguistic comparison. Gick et al. (2004) examined X-ray films of the vocal apparatus during ISP, at the midpoint of inter-utterance pauses. In Figure 2-1, adapted from Gick et al. (2004), an X-ray film frame corresponding to ISP is shown with indications of the measurements taken by the researchers. By comparing this figure with Figures 2-2 – 2-4, we see that the tongue in ISP is not in the position for the most open vowel, nor in the most raised position.

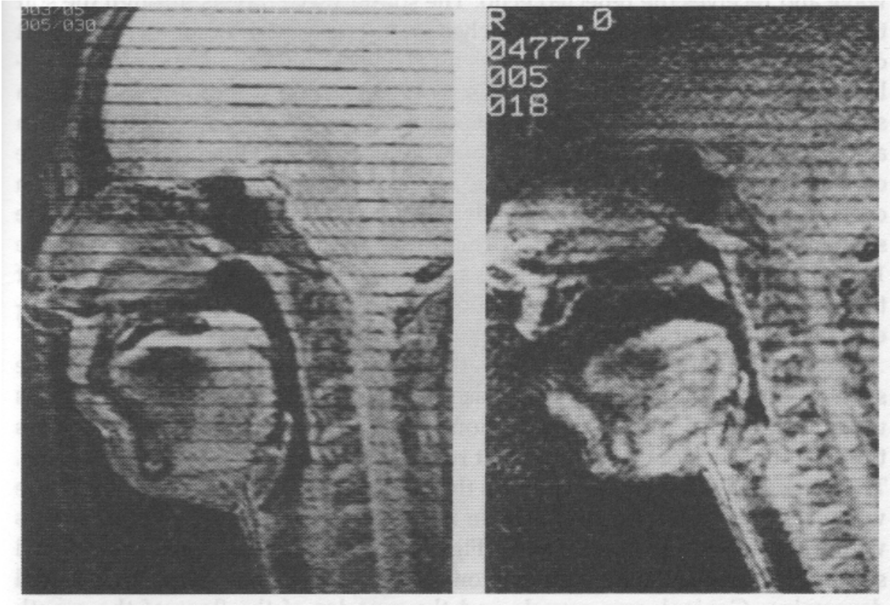


Measurements taken by Gick et al. (2004):

1 = pharynx width; 2 = velopharyngeal port width; 3 = tongue body distance from hard palate; 4 = tongue tip distance from alveolar ridge; 5 = lower-to-upper jaw distance; 6, 7 = upper and lower lip protrusion (distance from central incisors).

Figure 2-1. A frame of an X-ray film corresponding to the midpoint of an inter-utterance pause, vocal apparatus during inter-utterance speech rest position (after Gick et al. 2004).

When discussing the results of the analysis, Gick et al. (2004) say that the tongue shape for ISP may be close to the IPA schwa vowel’s articulatory configuration. However, these researchers have demonstrated that “schwa is not simply a vocalised instance of the ISP” (Gick et al. 2004, p. 231). For example, in English schwa, the tongue root retracts beyond the inter-speech posture (cf. Davidson and Stone 2003, who claim that the articulators’ position for schwa has a particular gestural target, referring to Browman and Goldstein 1992b).



a) b)

Figure 2-2. Two Magnetic Resonance Imaging (MRI) snapshots: a) vocal tract at rest; b) vocal tract during production of an /a/ (after Stone 1999). Note the similarity of the resting position with Figure 2-1.

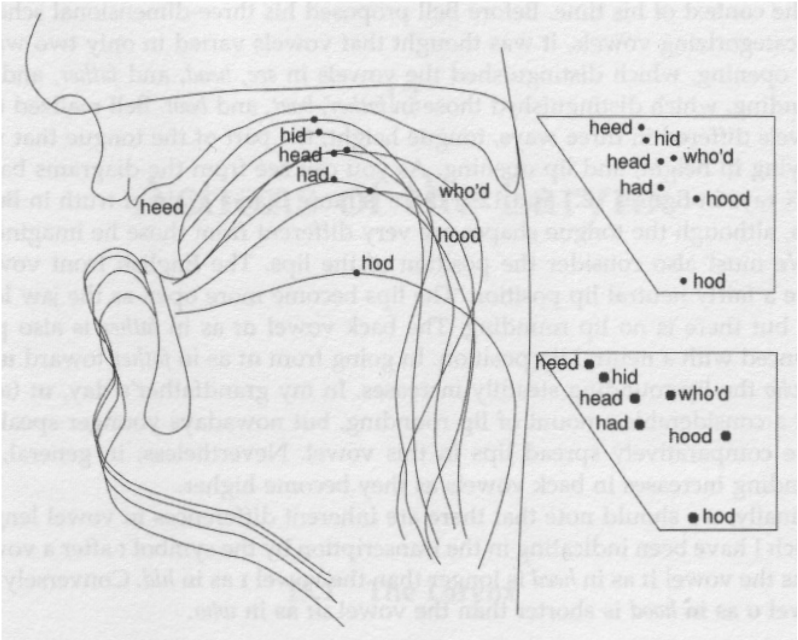
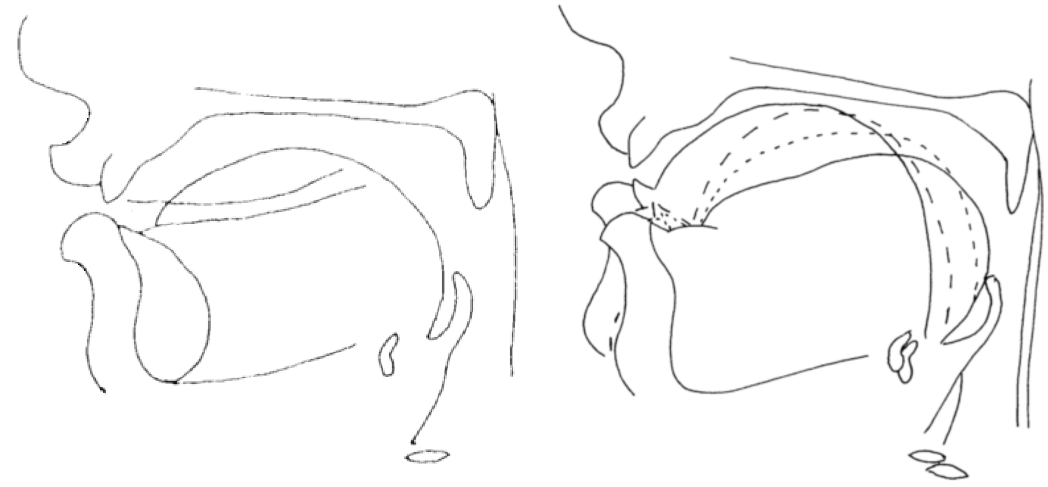


Figure 2-3. A diagram of the General American vowels based on X-ray data (after Ladefoged 2005). Note that the mid-closed vowel tongue positions are more similar to the resting position in Figure 2-1 and Figure 2-2a than the closed and the open vowel tongue positions.



a) b)

Figure 2-4. Midsagittal view of the vocal tract during production of a) /ə/ from the word /ʃəvalje/; b) /i/ (upper solid line), /e/ (dashed line), /ɛ/ (dotted line), /a/ (lower solid line). Based on French X-ray data. After Bothorel et al. (1986).

It has been shown that ISP varies across languages (e.g., Gick et al. 2004; Wilson 2005), and those differences can be explained by differing phonological systems. For example, Gick et al. (2004) observed the following differences in inter-speech posture between Canadian speakers of English and French: the tongue root was more retracted in English than in French, and the lower lip was more protruded in French than in English. These differences may have been due to the fact that there are more front vowels in French, and that French has an opposition with respect to rounding, in vowels. So it seems that even the "neutral" position is language-specific, and deactivation likewise.

2.2.3. "Syllabic" explanation of troughs

Thomas Gay was one of the first authors to investigate the physiological processes accompanying production of symmetrical and non-symmetrical vowel-consonant-vowel utterances (e.g., Gay 1974; Gay et al. 1974; Gay 1977). Gay (1974) noticed a cessation of activity for the genioglossus muscle during the time of consonant production. He concluded that each vowel in the sequence was marked by a separate muscle pulse. Gay interpreted his findings as contradictory to the two influential coarticulatory theories of the time. One of them was the look-ahead model (described below, Section 2.3.1), the other one was the numerical model of coarticulation (described below, Section 2.3.2.2). Gay et al. (1974) compared /VpV/ sequences in fast and slow speaking rates, and found that between the vowels, “the activity levels of the genioglossus muscle decrease during faster speech” (Gay et al. 1974, p. 53). In the study by Gay (1977), based on X-ray film data with the vowels /i, a, u/ and the consonants /p, t, k/, the following claim was made: “the finding that anticipatory movements begin and primary carryover effects end at about the same time during the closure period of the consonant, suggests that the release of the consonant and movement toward the vowel are organised and produced as an integral articulatory event” (Gay 1977, p. 192). Gay concluded that “the segmental input to the speech string is governed primarily by simple rules which act upon syllable sized units, while the temporal formulation of the string requires complex articulatory adjustments based on advance information obtained from a higher level scan-ahead mechanism” (Gay 1977, p. 192). These conclusions are reminiscent of the model presented in Kozhevnikov and Chistovich (1965), which states that the CV-type syllable is a unit of rhythm and a unit of articulation. For more details on syllable-based theories of speech production, including Kozhevnikov and Chistovich (1965), see Section 2.3.3.

2.2.4. “Syllabic” versus aerodynamic explanation of troughs

In Engstrand (1988), articulatory activity in Swedish symmetrical and non-symmetrical vowel-consonant-vowel sequences with the vowels /i/, /u/ and /a/ and the consonant /p/ was studied, using X-ray cinefilm and acoustic data. The purpose of the study was to measure differences in articulation between slow and fast speech rates. The trough pattern was observed in both subjects who took part in the study, and it was much more noticeable in the slow speech rate than in the fast speech rate. Troughs were evidenced by curvilinear patterns of the tongue movement between the vowels, “approximating a neutral position during the closure interval” (Engstrand 1988, p. 1867).

The author called the observed pattern “a relaxation of vowel-related tongue position”, and offered possible explanations. One of them drew on the model presented in Kozhevnikov and Chistovich (1965): “the articulatory relaxation observed here to precede the intervocalic consonant in /VpV/ utterances would be explained as a motor programming discontinuity at the syllable boundary” (Engstrand 1988, p. 1870). Another explanation suggested by Engstrand was actually very close to the first one. It was based on the idea expressed in Gay (1978) that articulatory relaxation in VCV sequences with symmetrical vowels and bilabial consonants is an expected pattern, because the CV portion constitutes a coarticulatory domain; hence, the vocal tract musculature would relax at the V/CV boundary, and reactivate when the articulatory programme enters the new domain.

Another of Engstrand’s suggestions of a possible reason for the trough pattern was the vocal tract shape for the production of the Swedish aspirated /p/. Referring to Stevens (1971), Engstrand hypothesised that “aerodynamic requirements on the production of the stressed /p/ release would include a relatively wide vocal tract. This condition could be met by delaying the vowel-related tongue movement as evidenced by the present subjects” (Engstrand 1988, p. 1871).

Engstrand also suggested that because of the aerodynamic requirement on the realisation of /p/, there may be differences in trough sizes in voiced versus voiceless consonants. In that study, the data for voiced consonants were not collected. The results of a later study by the same author on tongue behaviour in /p/ versus /b/ will be presented in Section 2.2.6.

2.2.5. “Diphthongal” explanation of troughs

In Engstrand (1988) described above, there was one more explanation of the trough patterns observed in the study, namely diphthongisation of stressed vowels in the dialect spoken by both experimental subjects. Engstrand hypothesised that “the presence and timing of the observed preclosure movement pattern may be understood without reference to notions such as ‘syllable’ or ‘segment’, but rather as a straightforward consequence of the speakers’ dialect” (Engstrand 1988, p. 1871).

Perkell (1986) is another published study of articulator deactivation in VCV sequences, which also mentions diphthongisation of the vowels surrounding the consonant as one of the factors possibly partially responsible for the deactivation pattern. In that study, Perkell did not analyse the tongue behaviour in VCVs with bilabial consonants, but rather the activity of the lips during /uCu/ sequences. He found "diminution of lip protrusion for the consonant in /uCu/ utterances" (Perkell 1986, p. 48), and he called that phenomenon a "trough". The researcher suggested that "the speaker might retract (either actively or passively) the lips or lower the tongue during the consonant in order to diphthongise the second vowel as well as the first" (Perkell 1986, p. 48). Perkell found differences in trough depths "which might reasonably be explained in part by 'underlying', language-specific differences in patterns of diphthongisation. However, there was no evidence of diphthongisation in the acoustic signal" (Perkell 1986, p. 64). Further on in the article, while naming possible mechanisms responsible for the production of troughs, Perkell mentioned "'underlying' factors which are not necessarily expressed in the acoustic signal", and among them "underlying language and/or individual-specific patterns of diphthongisation (which are not always manifested in the acoustics)" (Perkell 1986, p. 65).

2.2.6. Differences in trough patterns in voiced versus voiceless consonants: explanation by aspiration

After the publication of Engstrand (1988), a series of experiments was conducted by Engstrand and his co-author McAllister, in order to verify whether there was a difference in the trough patterns between voiced and voiceless consonants. In this series of experiments, intervocalic tongue relaxation in the Swedish symmetrical VCV utterances with voiced and voiceless bilabial stops was shown to differ (e.g., Engstrand 1989; McAllister & Engstrand 1991; McAllister & Engstrand 1992). For an explanation of the patterns found in the data, the authors adopted the hypothesis that aerodynamics of consonant production may account for the voicing variation in the lingual coarticulation discontinuity. One of the findings that made them think of such a hypothesis was that the trough had a greater magnitude for /p/ than for /b/. According to the argumentation of the researchers, these data could be explained by the fact that Swedish /p/ and /b/ have different aerodynamic conditions of production (crucially, the Swedish /p/ is aspirated): “the aerodynamic requirements on the production of the stressed, aspirated /p/ release would include a relatively wide vocal tract..., a condition met when the high vowel position is temporarily relaxed” (McAllister & Engstrand 1991, p. 20). So the researchers suggest that the need to lower the tongue during aspirated voiceless stops for precluding possible frication may be an explanation of the trough.

An argument against the explanation of trough occurrence by aspiration-induced acoustic or aerodynamic constraints was suggested in Lindblom et al. (2002). Lindblom and his colleagues refer to some of Engstrand’s unpublished data, and motivate their claim by the fact that in more recent data unaspirated and aspirated Swedish stops reveal identical trough patterns.

2.2.7. Differences in trough patterns in voiced versus voiceless consonants: explanation by tongue stiffening in /p/ versus tongue relaxation in /b/

Another aerodynamic explanation of coarticulatory patterns in VCV sequences was proposed in Svirsky et al. (1997). The aim of that study was to measure tongue surface displacement during bilabial stops, in order to gain more information about vocal tract impedance, “to test the competing claims of passive and active enlargement of the vocal tract during voicing” (Svirsky et al. 1997, p. 562). Tongue dorsum displacement perpendicular to the occlusal plane was measured with EMA. Svirsky and his colleagues observed tongue lowering in /aCa/ sequences with bilabial stops in English. The tongue displacement was smaller for /p/ than for /b/. The researchers explained the data they obtained by “active stiffening of the tongue during /p/, and/or... intentional relaxation of tongue muscles during /b/ (in conjunction with active tongue displacement during /b/), in order to accommodate airflow into the oral cavity while maintaining a transglottal pressure differential that will allow vocal fold vibration” (Svirsky et al. 1997, p. 570).

2.2.8. Aerodynamic explanation of troughs: intraoral pressure versus recombination of tongue and jaw movements

Explaining trough patterns in tongue movement in VCV sequences by aerodynamic constraints has been rather wide-spread among the studies concerned with troughs, as shown in Sections 2.2.4, 2.2.6 and 2.2.7. One more aerodynamic explanation is that troughs are due to the raised intra-oral pressure which occurs during oral stops pushing down on the tongue surface and some other requirement for the articulators to move to accommodate raised oral pressure. This explanation was taken up, for example, in Fuchs et al. (2004). That work used electromyography (EMG) and EMA. The researchers studied trough production in German VCV sequences with the consonants /p/, /b/ and /m/, and the vowels /i/, /u/ and /a/. The study was concentrated on verifying whether intervocalic tongue lowering may be due solely to aerodynamic constraints on producing different types of bilabial stops. The idea of the authors was that if aerodynamics alone is responsible for trough production, then there would be no need for the tongue lowering during /m/, because during production of nasals intraoral pressure does not rise as during voiced and voiceless stops.

Fuchs et al. (2004) found intervocalic tongue relaxation in their data (more in /p/, less in /b/, and much less in /m/). They also found that the extent of tongue lowering differed across muscles. Fuchs and her colleagues wrote that “if muscle fibers/tongue position can be modified by an increase of intraoral pressure…, known for voiceless stop production, our findings can be partially interpreted with respect to aerodynamic constraints” (Fuchs et al. 2004).

Fuchs and her co-authors found that the data showed trough patterns not only in VpV and VbV sequences, but in VmV as well. The authors concluded that aerodynamic requirements do play a role, but they alone do not cause troughs.

Also, in an experiment involving bite block, Fuchs et al. (2004) found support for the suggestion expressed by Vilain and her colleagues (e.g., Vilain et al. 1999), that troughs may be due to the recombination of tongue and jaw movements from V to C, and compensating by the new tongue position for the high position of the jaw implied by the lip occlusion. Cineradiographic data presented in Vilain et al. (1999) show troughs in symmetrical VCVs with bilabial consonants. For example, the researchers described the observed movements of articulators in /aCa/ sequences as follows: “the very low jaw height necessary for the production of [a] is brought back to zero for the consonant; yet the body of the tongue is not passively raised by this movement, as could be supposed. Instead we observe a reorganisation of the articulators, whose combined actions reconstruct the vocal tract shape of [a], without the contribution of the jaw” (Vilain et al. 1999, p. 170).

2.2.9. Explanation of troughs by the segment-by-segment activation pattern

The work of Björn Lindblom and his colleagues (Lindblom et al. 2002) was focused on studying troughs, and the term “trough effect” was used. The researchers aimed at “generating a more comprehensive characterization of the effect that can aid our understanding of its theoretical implication to speech motor control” (Lindblom et al. 2002, p. 246). The authors obtained acoustic evidence of trough production (the data were /ipi/ and /ibi/ tokens produced by five male speakers of American English). Transition durations of F2 into the labial closure extending beyond 10 ms were considered to have been contributed by slower tongue lowering/elevation movements, as compared with the transitions completed within a 10-ms interval, implying only labial closing/opening gestures. The authors also report the results of their analysis of some X-ray recordings of VCV sequences with /b/ and /p/. The data from one speaker (/ibi/ and /ipi/ sequences) demonstrate the trough pattern, both during /b/ and /p/, but somewhat larger and longer-lasting during the /p/ closure. The researchers use the term “deactivation” in the explanation of the observed patterns, as mentioned in Section 2.2.2. They suggest that muscle activity for the vowel is “turned off”, or deactivated, during the consonant, and activated again for the second vowel. A number of published studies are discussed in Lindblom et al. (2002). For example, the researchers discuss and compare Öhman’s numerical model of coarticulation (see Section 2.3.2.2) and the coproduction theory (see Section 2.3.2.3), and come to a conclusion in favour of the latter model: “speech motor planning can be more successfully modeled as a phoneme-by-phoneme event rather than diphthongal-like vowel trajectories with consonantal suppositions” (Lindblom et al. 2002, p. 250). The authors provide more support for the segment-by-segment activation pattern as the explanation for the trough effect from their vocal tract modelling data. However, the researchers note that the coproduction theory “falls short of providing a true explanatory account of the trough effect. A satisfactory account minimally needs to provide an explicit quantitative formalisation of the phenomenon” (Lindblom et al. 2002, p. 250).

2.2.10. More on the explanation of troughs by the segment-by-segment activation pattern; differences in tongue behaviour between high and low vowel contexts

The study by Vazquez Alvarez et al. (2004), and its follow-up study by Hewlett et al. (2004), aimed to discover whether or to what extent the trough pattern occurs during a bilabial consonant in symmetrical VCV sequences in British English. The data were the consonants /p/ or /b/, and the vowels /i/, /u/ or /a/. There were ten subjects. The data showed a strong tendency for a trough to occur in the two high vowels, though less in some subjects than in others. The authors suggested that the trough pattern was evidence of tongue deactivation (returning to the neutral position) during the bilabial closure, and that it supported the hypothesis by Lindblom et al. (2002) about a segment-by-segment activation pattern.

In /aCa/ sequences, there was a much lower incidence and extent of troughs. Not only intervocalic tongue lowering, but also tongue raising was observed. The numbers of occurrence of intervocalic tongue lowering and tongue raising patterns were nearly similar in /aCa/ sequences in that study. Note that these results on the coarticulatory patterns in /aCa/ sequences differ from those presented in Svirsky et al. (1997) and Fuchs et al. (2004): those researchers had observed tongue lowering between the two /a/ vowels in the VCV sequence, while Hewlett and his colleagues have documented the tendency to tongue raising between the two /a/ vowels. On the one hand, this was suggested by the authors to be the evidence of tongue deactivation (or returning to the neutral position) during the bilabial closure. On the other hand, the researchers argued that the upward movement of the tongue in /aCa/ sequences could be accounted for by the upward jaw motion required for the bilabial consonant production, and not necessarily by the need for tongue deactivation.

In the present study, /aCa/ sequences with bilabial consonants will be studied (Chapter 4), and the results will be discussed in relation to the findings presented in this section.

2.2.11. "Specification" explanation of troughs

The explanation of trough production by an underlying specification of the consonant has been proposed by some researchers. It has traditionally been thought that bilabial consonants are not specified for a particular position of the tongue, and that in languages which do not have a phonological opposition of labialised - nonlabialised consonants, consonants in general, unlike vowels, are not specified for the degree of lip protrusion. Thus, the look-ahead model (see Section 2.3.1) predicts that the tongue body should maintain the same position throughout the VCV sequence where the vowels are identical, and the consonant is a bilabial stop; and the degree of lip protrusion, according to the look-ahead model, should stay the same throughout an /uCu/ utterance. However, the experimental demonstration of the tongue/lip deactivation has shown that there are articulatory changes during these sequences. These findings have been interpreted so as to suggest that the consonants in question cannot be considered truly unspecified. There have been proposed a number of ways to account for this unexpected "specifiedness" of the segments previously thought to be unspecified. For example, Perkell (1986), having found troughs in /uCu/ utterances, suggested that alveolar consonants are specified for lip protrusion (see also Perkell and Matthies 1992). The researcher explained the production of troughs in labial protrusion in /uCu/ utterances by a few factors, and one of them was "the existence of specific lip protrusion targets for alveolar consonants which have previously been assumed to be 'neutral' for lip protrusion" (Perkell 1986, p. 66). The lip protrusion targets for the consonant /s/ varied across speakers and across languages (American English, Spanish and French were studied). For example, the lip protrusion target for /s/ in a speaker of American English was "somewhat between the targets for /a/ and /u/" (Perkell 1986, p. 63).

The concept of specification was used in some recent studies focused on studying troughs. For example, in a study by Fuchs et al. (2004), the authors, having observed a greater degree of trough occurrence in VCV sequences with /p/, /b/ than with /m/, suggest that labial consonants “may involve quite specific lingual adjustments” (Fuchs et al. 2004, C-30).

A critique of the explanation of lingual troughs in VCV sequences by tongue position specification was offered in Bell-Berti and Harris (1981). In that study, intervocalic suppression of EMG activity of the genioglossus muscle was observed. The researchers commented on the reasons as follows: “while it might be possible to claim that the consonant has some characteristic not actualised in its conventional feature description that causes the discontinuity, there is something unattractive about such an ad hoc solution” (Bell-Berti & Harris 1981, p. 12). However, Bell-Berti and Harris did not suggest why the muscle activity deactivated between the vowels.

2.2.12. Language-specificity in the production of troughs

The trough pattern in VCV sequences has been reported in some published works to differ across languages: it has been observed in some languages, but not in others. For example, McAllister and Engstrand (1991) studied the pronunciation of VCV sequences in one speaker of Swedish, one speaker of Australian English and one speaker of French, and found a consistent difference between French, on the one hand, and Swedish and English, on the other. In English and Swedish, where voiceless stops are aspirated, the relaxation of the tongue position coinciding with the consonant was statistically significant, and in French, where both voiced and voiceless stops are unaspirated, it was not. Moreover, the researchers found a difference in the tongue relaxation pattern between Swedish and English: the maximum of the relaxation occurred at the middle of the consonant in Swedish, and at the C/V2 boundary in Australian English. The authors ascribed this difference to a different relative timing of tongue articulation in these two languages.

Boyce (1990) studied similarly structured nonsense words including VCV sequences, in four speakers of American English and four speakers of Turkish. She found troughs in lip protrusion in English, but not in Turkish, a language which has vowel harmony. Boyce suggested that English and Turkish “may have different modes of coarticulatory organization” (Boyce 1990, p. 2584), and that the absence of troughs in Turkish was due to the language-related constraint being stronger than the articulatory tendency to deactivate the lip protrusion during the intervocalic consonant.

Engstrand et al. (1996) compared VCV utterances with bilabial oral stops in five Swedish speakers, three English speakers, two German speakers and three French speakers. Their EPG data showed troughs in Swedish and English, but not in German and French. Also, Swedish and English trough patterns were differently timed.

Some critical arguments about the language-specificity of trough production come from the study by Lindblom et al. (2002). The authors take a cautious approach to the idea that languages have different modes of articulatory organisation. They admit that “the trough effect appears to exhibit language-specific effects” (Lindblom et al. 2002, p. 261), referring to Perkell (1986) and McAllister and Engstrand (1992). But Lindblom and his colleagues also say that “departures from a general mode of coarticulatory organization across languages can be due to many, as yet unknown, phonetic factors, particularly as they relate to motor control constraints” (Lindblom et al. 2002, p. 248). While admittedly, some of the studies described in this section had a very limited number of participants, it nevertheless seems not accidental that cross-linguistic differences have been found for a number of languages, in independent studies.

2.2.13. Coarticulatory characteristics of [h]

We have seen in the previous sections that discontinuity in lingual coarticulation occurs between the two vowels in VCV sequences with bilabial consonants. In order to find out more about the mechanism of tongue behaviour during non-lingual consonants, let us look at the consonant [h]. This consonant, like bilabial stops, does not require any particular tongue constriction for its production, differing from the tongue configuration of the surrounding vowels. A difference between [h] and bilabial consonants is that [h] is not specified for lip and jaw position, while for producing labial consonants, both these articulators are constrained. Possible influence of the jaw and the lips on tongue position between the vowels of a VC sequence has to be accounted for when interpreting the results of the VCV coarticulation experiments. This has been mentioned by, e.g., Fuchs et al. (2004) and Vazquez Alvarez et al. (2004).

It is known from the literature that "[h] usually denotes a voiceless transition into (or, in some languages, out of) a syllable. Its place of articulation depends on the adjacent sounds" (Ladefoged 2001, p. 254). Keating et al. (1994) demonstrated that /h/ in English and in Swedish shows the greatest effect of vowel context, having compared symmetrical VCV sequences with the consonants /f, b, t, d, s, n, l, r, k, h/, and the vowels /i, e, a/. In their results, the jaw height range for /h/ was "much greater than that for any other consonant" (Keating et al. 1994, p. 415).

Several authors have argued that the production of [h] makes no demands on the supralaryngeal articulators, and, thus, [h] can be considered unspecified for the positions of supralaryngeal articulators (e.g., Keating 1988; Pierrehumbert & Talkin 1992). Boersma (1998) writes that [h] "violates the complementarity of sonorants and obstruents, since it is not a sonorant (i.e., there is no perception of voicing) and it is not an obstruent either (i.e., there is no strong supralaryngeal constriction)" (Boersma 1998, p. 17). Boersma considers [h] to be placed together with voiceless fricatives on the sonority scale: "phonetically, it is a voiceless fricative whose noise stems from the glottal constriction and from any other places in the vocal tract that happen to be narrowed; though its spectral properties depend strongly on the shape of the supralaryngeal cavities, we would be inclined to classify it with the low-sonority voiceless fricatives /fsx/ in the hierarchy" (Boersma 1998, p. 16).

The claim about [h] being unspecified for the positions of supralaryngeal articulators has been experimentally supported by ultrasound. Karbownicki (2004) studied coarticulation in British English /hV/ syllables, both words and non-words, aiming "to investigate coarticulation effects on the glottal fricative [h] when it precedes a vowel" (Karbownicki 2004, p. 14). The study tested the claim by Ladefoged (2001) that [h] is merely a voiceless version of the following vowel. The results of the study showed that the tongue shape during [h] was very similar to the tongue shape during the following vowel, across several vowel contexts. The author interpreted these results as strongly supporting Ladefoged's (2001) statement about [h]'s lack of its own place of articulation.

2.3. Theories of speech production relevant for this research

In this section, several theories of speech production are described. Concepts or claims from these theories are relevant for the present study, because they set the general context for working on coarticulation in speech. In describing the theories, the accent is on how each of them treats coarticulation in VCV sequences.

2.3.1. Feature-based models

Some concepts of feature-based theories of speech production are close to the notions used in the CR approach to coarticulation. For example, specification of a speech sound for a particular tongue position within the CR approach implies that tongue position is linguistically relevant for this sound. In other words, changing the tongue position may result in a phonological contrast. Thus, the consonants /t/ and /k/ can be described as forming an opposition, based on the tongue posture during the consonant occlusion and burst. So the concept of phonological opposition, as described in Trubetzkoy (1939), is highly relevant for the CR approach to speech production.

Another concept coming from early feature-based phonological works is markedness. The concept of markedness was first introduced into phonology and linguistics by the Prague School, particularly developed by N. Trubetzkoy in his Gründzuge der Phonologie (Trubetzkoy 1939), and later it was widely adopted by linguists. In the Oxford Encyclopedia of Language and Linguistics, D. Archangeli links markedness in phonology "with the representation of the asymmetric distribution of segments in sound systems. Where an opposition is possible, the typical pattern or property is called UNMARKED, and the atypical one MARKED" (Archangeli 1992, p. 391).

The term "specification" also comes from feature-based phonological studies. When characterising the system of phonemes within distinctive feature frameworks (e.g., Jakobson et al. 1951; Chomsky & Halle 1968), each phoneme is described in terms of a set of features, forming oppositions. For example, in Jakobson et al. (1951), the system of twelve binary acoustic distinctive features was used, and the authors aimed to introduce a system that would be able to describe phonetic inventories of all human languages. The system of binary distinctive features implies that the phonemes of the language are characterised by a number of features, which can have either of the two values: [+] or [–]. Some phonemes may be unspecified for some features, i.e., these features do not oppose these phonemes to others. For example, if we have to characterise a bilabial oral stop (in a system with three labial consonants: /p/, /b/, and /m/), then the linguistically relevant features will concern lips (for distinguishing this phoneme's place of articulation from other phonemes' places of articulation), velum (oral/nasal), and larynx (voiced/voiceless). We do not need to specify anything about the tongue, as its position is not linguistically relevant for a bilabial sound.

An articulatory model of speech production based on distinctive features and oppositions was presented in Henke (1966), and later became known as the "look-ahead" model. According to this model, the input to the articulatory system is a string of phoneme-sized units. These units are specified by a bundle of component features. According to the look-ahead model, if a particular segment in a string of segments is specified for a certain feature, this feature is automatically assigned to all the preceding segments unspecified for this feature. The "look-ahead" mechanism scans strings of segments, and assigns "upcoming" features to the segments as early as possible, so long as the anticipated features do not enter in contradiction with the articulatory requirements on the preceding segments. For example, the look-ahead model predicts maximum vowel-to-vowel lingual coarticulation across a bilabial stop, because the tongue is not required to have any particular position for producing bilabial stop consonants.

The principles of operation of the look-ahead mechanism are exemplified in Figure 2-5. In the figure, the assignment of feature values for upper lip protrusion in a French utterance is illustrated, based on the data from the study by Benguerel and Cowan (1974). Line 1 is the phoneme sequence input to the model. Line 2 specifies the category for each segment. There are four categories: "V_u" – unrounded vowel, "V_r" – rounded vowel, “C” – nonlabial consonant, “C_l” – labial consonant. The symbol “#” signifies the end of utterance. In line 3, the upper lip protrusion feature value at the phonemic level is indicated. There are three posible values, and the segment categories have the following specifications: unrounded vowels are specified for the feature value “–”, rounded vowels and labial consonants are specified for the feature value “+”, and nonlabial consonants are unspecified for lip protrusion (feature value “0”). In line 4, feature specification at the articulatory goal level is presented. It demonstrates that all the segments unspecified for lip protrusion at the phonological level (as shown by zeros in line 3) receive the feature value “+” at the articulatory goal level, if they occur before a segment having the feature value “+” at the phonological level. So in the phoneme sequence “strstrykty”, all the phonemes have a positive specification for lip protrusion at the articulatory goal level. Line 5 contains a graphical representation of the feature specification at the articulatory goal level, and line 6 is a “coarse approximation to the continuous output of the model” (Benguerel & Cowan 1974, p. 52).

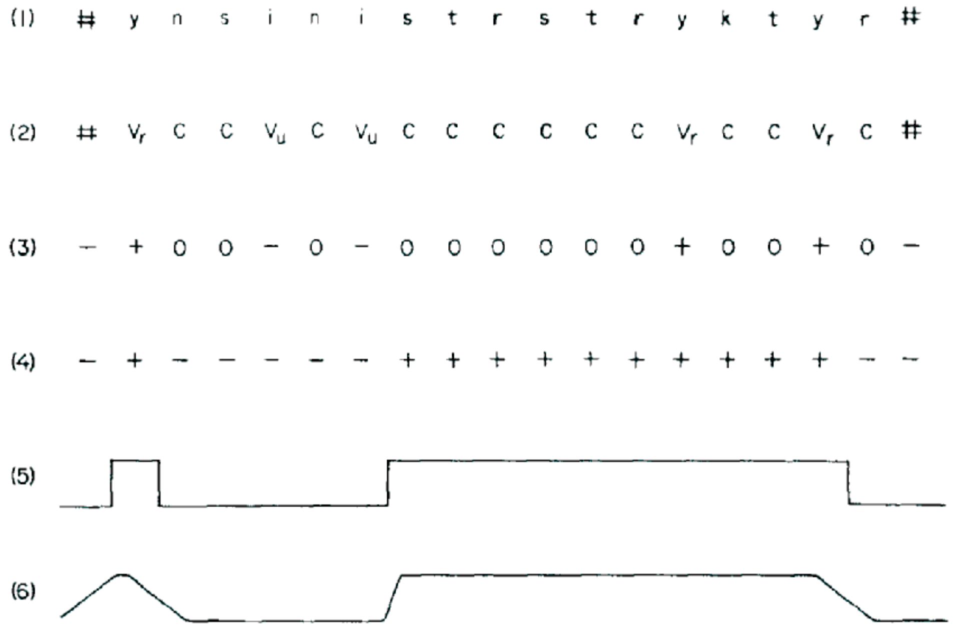


Figure 2-5. A diagram illustrating the prediction of the protrusion gesture time sequence for the French utterance “une sinistre structure” (after Benguerel and Cowan 1974).

Experimental support for the look-ahead model also comes from, for example, Moll and Daniloff (1971). These researchers observed anticipatory velar coarticulation in vowels preceding nasal consonants, and explained their findings within the framework of the look-ahead model. For example, the velar coarticulation pattern in CVVN sequences (where N stands for a nasal consonant) was explained as follows: "for a CVVN sequence, it would be predicted that velar opening for the nasal consonant would be initiated at the beginning of the first vowel in the sequence, a prediction in agreement with the data obtained in this study" (Moll & Daniloff 1971, p. 683).

Some general criticism of binary feature coding models was formulated, for example, in Kent (1983). The researcher claimed that these models "do not satisfactorily predict the finer variations of articulatory timing. Because any predictions of movement timing are tied to a segmental input by means of the intervening feature coding, coarticulation occurs only as the result of feature spreading across segments" (Kent 1983, p. 63). The implication here is that fine articulatory changes not coinciding with segment boundaries cannot be predicted by segment-based models.

2.3.2. Gesture-based models

After the appearance of the look-ahead model, it was shown by different authors that the model fails to account for some coarticulatory phenomena present in natural speech (e.g., Bell-Berti & Harris 1974; Gay 1978; Bell-Berti & Harris 1981; Sussman & Westbury 1981; Lubker & Gay 1982). These failures largely concern the static, non-flexible predictions of the model in terms of segments and features. The idea of speech units as dynamic gestures became popular in phonetic studies approximately at the time when the facts contradicting the look-ahead model were presented. Representation of the units of speech as being inherently dynamic gives the researchers flexibility in modelling human speech, which is continuous by nature, and the elements of which co-occur and interact. In this section, some of the dynamically oriented models will be described, in relation to the topic of this study.

2.3.2.1 Temporal model of speech production (Bell-Berti & Harris)

Some experimental data contradicting the predictions of the look-ahead model became a starting point for the model described in this section.

One of the experimental findings was that in a string of unspecified segments preceding a segment specified for a particular feature, the onset of an articulator's movement towards a specified position does not necessarily occur during the first segment in the string, but rather it is locked to a fixed time before the specified segment. A number of studies have observed that the anticipatory muscle activity for an upcoming target is beginning at a fixed time before the target realisation, regardless of the number and duration of preceding segments: e.g., Bell-Berti and Harris (1979); Bell-Berti and Harris (1981); Bell-Berti and Harris (1982). For example, Bell-Berti and Harris compared onset times for EMG activity from the orbicularis oris muscle in the sequences /i#tu/, /is#tu/, /i#stu/ and /is#stu/ ("#" meaning word boundary), and the results demonstrated that EMG activity begins "a constant time before the beginning of the acoustic period for the vowel" (Bell-Berti & Harris 1981, p. 14).

Another finding contradicting the predictions of the look-ahead model was that the tongue gesture for the vowels in the sequence [ipi] was demonstrated to be not continuous; instead, a cessation of genioglossus muscle activity was observed between the vowels (e.g., Bell-Berti & Harris 1974). This fact could not be explained by the look-ahead model, which would predict that "the tongue gestures for the vowels in the sequence [ipi] should be continuous, since there is nothing in the unrounded vowel gesture which conflicts with the production of the labial consonant" (Bell-Berti & Harris 1981, p. 12).

A model that was presented in Bell-Berti and Harris (1981) appeared out of experimental work on coarticulation, largely using VCV sequences as the data. The proponents of the temporal model of speech production claim, based on their own findings (e.g., Bell-Berti & Harris 1974; Bell-Berti & Harris 1979) and other reports in the literature (e.g., Gay 1978), that timing is "an integral organizing parameter of the speech motor plan", that time and timing relationships are "intrinsic to speech motor organization", and that the units of speech are "inherently dynamic gestures rather than static vocal tract configurations or invariant commands to the articulators" (Bell-Berti & Harris 1981, p. 9).

Bell-Berti and Harris (1981) state that "the temporal relationships among the articulatory components of a particular segment are integral to speech organization" (Bell-Berti & Harris 1981, p. 9). The researchers then propose three rules for the timing of articulatory activity that underlies segment representation.

The first rule is that the articulatory period of a segment is longer than its acoustic period. The rule represents the idea that the movements toward an articulatory goal and away from it form part of the segment's specification. This rule is largely incorporated in the principles of articulatory phonology, as we will see later (Section 2.3.2.3).

Rule two is that "for a given articulator the period of anticipation is temporally independent of preceding phone string length, if there is no articulatory conflict" (Bell-Berti & Harris 1981, p. 14). The phenomenon described in this claim was later referred to as "time-locking" (e.g., Perkell 1986).

The third rule states that the articulatory period may begin at different times for different articulators. It follows from this rule that "the durations of the parts of the articulatory period preceding and following the acoustic period must be empirically determined for each articulator" (Bell-Berti & Harris 1981, p. 16).

Bell-Berti and Harris say: "we believe that phonological units are inherently dynamic and retain this essential property when they are produced… The common failure to uncover static units… does not indicate that speech production obscures the static properties of segments, but rather that the articulatory stream has been described in the wrong way" (Bell-Berti & Harris 1981, p. 16).

2.3.2.2 Numerical model of coarticulation (Öhman)

One of the influential coarticulatory theories in the 1960s was introduced by Öhman (Öhman 1966; Öhman 1967). The researcher studied the articulatory behaviour of vowel-consonant-vowel sequences with different vowel combinations and several consonants, in Swedish, American English and Russian. Based on the experimental results, Öhman developed the numerical model of coarticulation accounting for different coarticulatory patterns in these sequences (Öhman 1967). Öhman compared the sequences with labial, alveolar and velar stop consonants, and found that the values of F2 transitions between adjacent segments (V1-C and C-V2) depended to a certain extent on the quality of the other vowel (V2 and V1, respectively). These data were interpreted by Öhman as evidence that vowel-consonant-vowel sequences are produced as a vowel-to-vowel diphthongal gesture with a separate consonantal gesture superimposed on it. He wrote: "The vowel and consonant gestures are largely independent at the level of neural instructions" (Öhman 1967, p. 310).

There have been later attempts to experimentally investigate whether the underlying mechanism of coarticulation in vowel-consonant-vowel sequences was indeed working as predicted by Öhman's model. For example, Ericsdotter et al. (1999) attempted to account for vowel-consonant-vowel coarticulation in Swedish, using the model developed by Öhman. The researchers found that the place of consonant articulation "varied markedly as a function of the preceding vowel" (Ericsdotter et al. 1999, p. 1886), and that the consonant dependence on the vowel context was much stronger than could be accounted for by Öhman's model. The same group of researchers studied symmetrical /iCi/ sequences with bilabial consonants, and they observed a discontinuity in the F2 pattern, that lasted for longer than just a bilabial closure effect, and was suggested to be an effect of discontinuity in tongue movement between the vowels (Lindblom et al. 2002). Lindblom and his colleagues concluded that the observed pattern could not be predicted by the numerical model of coarticulation (for more on Lindblom et al., 2002, see Section 2.2.9).

There is an interesting statement from Öhman, that does not quite fit in with the interpretation made by Lindblom et al. (2002): "The production of the consonant involves concomitant articulatory adjustments partially anticipating the configuration of the succeeding vowel... furthermore, the medial consonant configuration may be slightly anticipated during the initial vowel" (Öhman 1966, p. 168). This claim made by Öhman allows for the vowel-to-vowel diphthongal gesture interpretation of his data, but does not necessarily require that the vowel-to-vowel transition cannot be influenced by the medial consonant. This is further supported by Öhman's data on vowel-consonant-vowel sequences in Russian. Unlike Swedish and American data, Russian vowel-consonant-vowel articulations did not exhibit significant variations of V1-C transitions as a function of the identity of V2. Öhman explained this by the fact that Russian consonants, which have to be either palatalised or velarised, imply special neural commands to the tongue, and thus the consonant gesture overrules the vowel gesture and coarticulation is blocked. In this work, one of the experiments is focused on VCV sequences with Russian bilabial consonants (see Chapter 4).

2.3.2.3 Articulatory phonology

The claim about the independence of vowel and consonant gestures made by Öhman (see Section 2.3.2.2) was taken up and developed in later studies. For example, it was shown in Perkell (1969) that consonants and vowels are produced by different sets of muscles. Fowler (1980) argued that these differences in the ways of producing vowels and consonants and in their spatiotemporal properties allow them to be coarticulated (see also Fowler et al. 1980). Fowler claimed that a plausible way of accounting for the apparent coproduction of vowels and consonants "is for the articulatory systems responsible for vowel and consonant production to be distinct" (Fowler 1980, p. 129), and that "the capacity for coproduction derives from an adaptive property of speech that the two classes of articulatory gestures, consonants, and vowels, are products of different (coordinated) neuromuscular systems" (Fowler 1980, p. 129). Starting from these claims, Fowler and her colleagues developed the coproduction theory of speech production, which became one of the foundation stones for articulatory phonology (see Section 2.3.2.3).

The coproduction theory, or the intrinsic timing model (e.g., Fowler 1977; Fowler 1980; Fowler et al. 1980), claims that the original phonological units are not abstract, static and timeless, as suggested by the featural theories, but instead they are dynamically specified phonetic gestures which have their own intrinsic temporal dimension. This idea is quite consistent with the temporal model of speech production, described in Section 2.3.2.1. Ostry et al. (1996) describe this approach to speech production as follows: "coarticulatory changes result from the temporal overlap of control signals associated with the production of vowels and consonants" (Ostry et al. 1996, p. 1571).

Phonetic gestures are also the core of the task dynamic model associated with the coproduction theory (e.g., Saltzman & Kelso 1987; Saltzman & Munhall 1989; Saltzman 1991; Saltzman & Byrd 2000). The task-dynamic model, developed for explaining non-speech movements, and later applied to speech, is described in Hawkins (1992) as "a particularly useful way of analyzing speech production, partly because it breaks complex movements down into a set of functionally independent tasks" (Hawkins 1992, p. 9). In the task-dynamic model, "articulatory movement patterns are conceived of as coordinated, goal-directed gestures that are dynamically defined… This approach captures coproduction by allowing gestures to overlap in time, with the acoustic consequences of the coproduced units reflecting their combined influence on the vocal tract" (Saltzman & Byrd 2000, p. 501).

Articulatory phonology is currently a very popular theory of speech production, that aims to both predict and model the articulatory behaviour of human beings. The proponents of the model call it a "phonology for public language use" (Goldstein & Fowler 2003). The theoretical framework of articulatory phonology, combining the principles of the coproduction theory and their application in the task dynamic model, is presented in, e.g., Browman and Goldstein (1990), Browman and Goldstein (1992a). The model has its roots in the research that was described in Sections 2.3.2.1 and 2.3.2.2. Researchers who work within the framework of articulatory phonology consider gestures, and not features, to be the underlying cognitive linguistic units. Fowler and Saltzman (1993) define phonetic gestures as "linguistically significant actions of structures of the vocal tract" (Fowler & Saltzman 1993, p. 172). Gestures are dynamic units, which, unlike features, are allowed to overlap in time. According to the theory, gestures as phonological units are actualised in speech, but they are not modified in the speech continuum, rather, due to their intrinsic temporal structure, they can overlap in time with neighbouring gestures. They are coproduced with the context, rather than modified by the context.

In the case of coarticulation in VCV sequences, according to articulatory phonology, the coproduced vocalic and consonantal constriction gestures involve certain articulators, and there occurs a gestural interference. The notion of gestural interference refers to the fact that when neighbouring gestures are coproduced in speech, the degree of coarticulation depends on whether the same articulators are shared by the overlapping gestures or not. For example, articulatory phonology claims that lingual coarticulation is maximal in VCV sequences with non-lingual consonants, as the articulators for vocalic and consonantal gestures are not shared, and the gestural interference is minimal.

Articulatory phonology could incorporate the findings described in Section 2.2, about the discontinuity in lingual coarticulation in VCV sequences, in two different ways. One of them would involve the concept of neutral position. Articulatory phonology claims that an articulator returns to a neutral position when it is not part of any active gesture. Each vocal tract variable has its neutral gesture. The role of this gesture in the task dynamic model "is to bring the vocal tract articulators back to a neutral position when they are not otherwise being actively controlled by a constriction gesture. (Without a neutral attractor, articulators could simply be 'stuck' in a constricted posture if not called away by another gesture.)" (Byrd & Saltzman 2003, p. 164). In the case of a lingual trough in symmetrical VCV sequences with bilabial consonants, articulatory phonology assumes that the tongue does not form part of an active gesture for the bilabial consonant, and hence it is deactivated between the two vowels. Another way of incorporating discontinuity in lingual coarticulation in VCV sequences with non-lingual consonants in articulatory phonology would be to assume that non-lingual consonants have a particular position of the tongue as part of their active gesture. So the model can incorporate the data in either of these two ways, and it is not clear which one of the two ways would better predict the coarticulatory pattern in a given language. The DAC model, as compared with articulatory phonology, only has one way of interpreting these coarticulatory patterns – by specification of tongue position. The DAC model does not incorporate the notion "deactivation". This forces the model to make more precise predictions than articulatory phonology.

Articulatory phonology incorporates various syllable-related coarticulatory phenomena. For representing syllable structure, the model has two different functional tiers, a vocalic one and a consonantal one. These two tiers can capture “the fact of articulatory overlap between the vowels and consonants” (Browman and Goldstein 1990, p. 352). “Association lines” between the two tiers indicate that a consonantal gesture co-occurs with a vowel gesture. The task dynamic model associated with the theory states that the difference between CV and VC sequences is in the phase coordination. There is the “in-phase” coordination between the consonant and vowel of the CV sequence, i.e., the gestures for the consonant and for the vowel begin at the same time. In the VC sequence, there is the “anti-phase” (180˚) coordination, responsible for the smaller extent of coarticulation in this sequence (e.g., Goldstein et al. 2006).

This description of articulatory phonology shows that it has many technical possibilities for accounting for coarticulatory phenomena in VCV sequences. In this work, the DAC model is used as a theoretical framework, and not articulatory phonology, because the former is more focused on lingual coarticulation, largely based on the data from VCV sequences, and offers more precise predictions than the latter, as shown in this section.

2.3.2.4 Window model of coarticulation

Another framework designed to analyse dynamically changing articulations is called the window model of coarticulation (e.g., Keating 1988; Keating 1990). In this model, “continuous spatial representations are derived using information about the contextual variability, or coarticulation, of each segment” (Keating 1990, p. 452).

According to the window model of coarticulation, a segment may be specified or unspecified at either the phonological representation level or the phonetic representation level, or both. If the segment is specified not only on the phonological, but also on the phonetic level, then its range of articulatory or acoustic variation may be kept small through the specification of a quite precise window. If the window is narrow it will not permit the articulation to show wide contextual variation. If the segment is left unspecified altogether, its window will be wide and will allow the segment to have very large contextual variation.

In the window model, the degree of contextual variation is language-specific and is defined by the size of the windows, which is derived from the information on contextual variability in speech. Phonetically, "a wide window specifies relatively little about a segment, while a narrow window gives a precise specification, and all intermediate degrees are possible. Thus, for example, with respect to phonetic nasality, English vowels, with their wide but not maximal window, are "not quite unspecified" (Keating 1990, p. 465). So the window model has a possibility for accounting for possible resistance to coarticulation in speech sounds by having various degrees of phonetic specification.

Large lingual coarticulation in non-lingual consonants, according to this model, results from the fact that the consonants in question are unspecified for lingual features in the phonological representation. For example, the consonant /h/ is described in the window model as having no oral features in its phonological representation. Lingual consonants have precise phonological specification for their lingual feature windows, and this results in less influence from the vowels on these consonants. Vowels, presumably, have even narrower windows for their phonological representation of lingual features.

In the window model of coarticulation, syllable structure is not explicitly incorporated into the coarticulatory scheme based on windows of different size. There is no information in the model on how syllable division present in speech is related to the size of the windows, and how it is related to the "contour building", or finding a path through the windows. Neither does the model describe in detail "how timing fits into this scheme, e.g., the time interval over which paths are constructed, and whether windows have variable durations, or are purely notional" (Keating 1990, p. 457).

The window model states that "a language can have wider or narrower windows for all feature values" (Keating 1990, p. 467). According to the model, there is "more than one possible path through a given sequence of windows, especially at the edges of the utterance. This indeterminancy is seen as an advantage of the model; it says that speakers truly have more than one way to say an utterance, especially in cases of minimal context" (Keating 1990, p. 462).

For interpreting lingual coarticulatory patterns in VCV sequences with non-lingual consonants, the window model could offer several ways. The intervocalic consonant may be considered specified for a particular tongue position, or else "not quite unspecified". Any of these possibilities would accommodate a continuous tongue movement between the vowels, but also a trough pattern. However, having three different possibilities of explaining a phenomenon, it is impossible to make precise predictions of what would happen. So the window model appears to be too vague for using as a framework for this research, as compared with the DAC model.

2.3.3. Syllable-based models

The model presented in Kozhevnikov and Chistovich (1965), later called the "syllable-based model of articulatory encoding" (Bladon & Al-Bamerni 1976), drew on extensive research on coarticulation in Russian. Kozhevnikov and Chistovich (1965) studied temporal and spatial aspects of coarticulation within a phrase, articulatory and acoustic characteristics of different types of syllables, and temporal relation of the sounds within a syllable. Based on their own findings and on previously published literature on anticipatory coarticulation (e.g., Halle et al. 1957; Ladefoged 1957), the researchers claimed that within a phrase, speech is organised into syllables of CV-type, all consonant sequences syllabify with the following vowel, and this articulatory syllable is the unit of speech encoding.

A complex methodology was used in Kozhevnikov and Chistovich (1965) for the combined acoustic and articulatory analysis of speech, and perceptual experiments were also conducted. The articulatory techniques included registering intra-oral pressure, oral and nasal airflow, EPG, combined video- and photo-oscillographic recording of the lower part of the face (including the jaw and the lips), pneumographic registration of breathing, registration of larynx activity using a microphone, and registration of lip movement by means of electroplethismography.

One of the interesting findings that made Kozhevnikov and Chistovich come to their conclusions was the following. They conducted an experiment on “artificial stuttering” in the condition of delayed auditory feedback. The subjects produced phrases that they had to learn before the experiment. The phrases were recorded onto audiotape, and played back to the subjects with a short delay (100, 210, 300 and 470 ms after production). The effect of the delayed auditory feedback on the production was studied articulatorily and acoustically. The most salient feature found by the authors was multiple repetitions of CV-type syllables. When the phrases contained consonant sequences, these sequences were broken into CV complexes, where V was a schwa-like vowel. The researchers experimented with different places of morphological boundaries, and different types of consonant sequences, but irrespective of these conditions, when auditory feedback was delayed the minimal unit produced by the speakers was a CV combination. These findings made the authors conclude that the basic speech units are CV-type articulatory complexes, and that more complex sequences, such as CCV or CCCV, are just combinations of these simple units, organised in such a way that the following complex starts earlier than the preceding one is finished.

Later experimental studies found evidence against the model proposed in Kozhevnikov and Chistovich (1965). For example, Moll and Daniloff (1971) observed anticipatory velar coarticulation starting at the beginning of the first vowel in the English sequence consonant-vowel-vowel-N, where “N” was a nasal consonant. The researchers concluded that their results directly contradicted the hypothesis that a CV-type syllable was the minimal unit of coarticulation and production. More evidence contradicting the model comes from Benguerel and Cowan (1974). They studied groups of VCC...CV sequences (with the first unrounded vowel, and the second rounded vowel), and found that the lip protrusion movement in French may start during the vowel preceding the consonant cluster. These results were interpreted as contradicting Kozhevnikov and Chistovich’s model of a CC...CV-type syllable.

Zinder (1979) criticised the syllable-based model as follows: “This answer to the question of syllable division in speech may describe the situation in Russian to some degree, but it cannot be accepted as universal” (Zinder 1979, p. 256). Zinder motivates his claim by referring to acoustic characteristics of German, where syllable division depends on whether the vowel preceding a consonant is long or short.

After Kozhevnikov and Chistovich (1965), many studies focused on analysing characteristics of CV versus VC syllables, and produced evidence of stronger coarticulation in CV than in VC syllables: e.g., Fromkin (1968), Bondarko (1969), Gay (1978), Perkell (1986), Browman and Goldstein (1988), Byrd (1995), Lindblom et al. (2002). For example, Browman and Goldstein (1988) studied the timing of consonant sequences with neighbouring (preceding and following) vowels. The researchers used the term “C-center”, to denote the centre of a sequence of consonant gestures. Postvocalic consonant sequences, according to Browman and Goldstein (1988), were timed to the preceding vowel locally by the left edge of the first consonant, regardless of the syllabic affiliation of this consonant. “Local timing” means articulatory coordination of a single consonant gesture in a consonant sequence with a neighbouring vocalic gesture. Prevocalic onset consonant sequences (forming one syllable with the following vowel), according to the authors’ claim, were timed to the following vowel globally, by their C-center. “Global timing” implies articulatory coordination of a consonant gesture sequence taken as a single unit with a neighbouring vocalic gesture. These results show that sequences “consonants – vowel” are more closely connected than sequences “vowel – consonants”, because all prevocalic consonants belonging to one syllable are timed to the following vowel, while only the first consonant of a postvocalic sequence is timed to the preceding vowel.

Syllables are considered to be basic segmental units in speech production in two current models of articulatory organisation of speech – the C/D model (e.g., Fujimura 2000), and the time structure model of the syllable (e.g., Xu & Liu, forthcoming). The C/D (“Converter/Distributor”) model is “a computational description of articulatory gesture organisation using syllables, rather than phonemes, as the concatenative segmental units” (Fujimura 2000, p. 129). The time structure model states that the syllable specifies the temporal alignment of the basic phonetic elements, including consonants and vowels, and that the “universal stability of the CV structure… is the natural outcome of the basic temporal alignment pattern of the syllable" (Xu & Liu, forthcoming).

Syllable-based models are not chosen as a framework for this research because they do not make specific claims about some questions addressed in this work, for example, how different vocalic contexts influence the same consonant, or whether the same vowel in different consonant contexts has significantly different tongue shapes.

2.4. Conclusions

The aim of this chapter was to present the theoretical framework for this research (the CR approach and the DAC model), to describe relevant previous studies of coarticulation in VCV sequences, and to explain the choice of this framework, by comparing it to other theories of speech production.

In this chapter, gesture-based and feature-based models have been described. Gesture-based models have been shown to better account for coarticulatory processes in speech, because of the inherently dynamic nature of gestures. However, an interesting point is that both gestures and features are based on phonemes of the language. For example, the DAC model claims that gestures are underlying units in speech production. But the unique characteristic of this model and the reason for choosing it as a framework for this research is the concept of CR.

Resistance to coarticulation is a concept that allows for describing in numbers how much speech sounds influence neighbouring sounds, and at the same time, how much they are influenced by neighbouring sounds. This is a nice philosophy for studying phonetics. The DAC model has a rather well-developed terminological apparatus for analysing coarticulation. As compared with other contemporary models of coarticulation (e.g., articulatory phonology, or the window model), the DAC model is narrow enough for making quite precise predictions about lingual coarticulatory behaviour, based on the notion of articulatory constraint. The existing publications within the DAC framework do not explicitly describe how cross-language and cross-subject variation would be incorporated in the model. However, the predecessors to the DAC model, Bladon and

Al-Bamerni (1976), do describe CR as being language-specific, and even subject-specific.

There are two more reasons for choosing the DAC model as a framework for this research. One is that the articulatory constraints on production of speech sounds are not yet formulated well enough to produce a non-contradictory description of some phenomena. For example, non-lingual consonants are treated by the model as being minimally resistant to lingual coarticulation. It is indeed quite well established in the literature that non-lingual consonants are less resistant to lingual coarticulation than the consonants involving the tongue in their production (e.g., Recasens 1999). However, there are no commonly accepted satisfactory accounts of how much "less resistant" they are. Within the CR framework, there are few studies on consonants with minimal CR. Results from different studies made not within the CR approach suggest that lingual behaviour in VCV sequences with non-lingual consonants may be language- or even subject-specific (e.g., Engstrand 1988; Lindblom et al. 2002).

Another attractive characteristic of the model is that it uses the same scale of articulatory constraints for consonants and vowels, based on the tongue dorsum involvement in their production. The relationship of vowels and consonants in speech can be expressed in numbers within this model. It would be interesting to investigate the difference in CR between groups of speech sounds with varying degrees of coarticulation resistance – consonants of different place and manner of articulation, and vowels.

Finally, the DAC model makes specific predictions which can be tested using ultrasound. Ultrasound captures the movements of the tongue body, and it can be safely used for obtaining as much data as needed. Theoretical assumptions made in the DAC model about lingual coarticulatory behaviour are mainly based on EPG studies of tongue movements. EPG, unlike ultrasound, is an indirect method of registering tongue movements. So it is very interesting to apply ultrasound in order to test the model's predictions. The next chapter is dedicated to the method used in this work.

The overarching goals of this research are to combine the DAC model as a theoretical framework and ultrasound as the principal method of investigation, and to study mutual coarticulatory influences of different groups of speech sounds. The issues to be addressed include vocalic influence on consonants, consonantal influence on vowels, and the role of a syllable boundary in coarticulatory interaction of speech segments. A further goal is to attempt to quantify lingual CR characteristics of vowels, lingual and non-lingual consonants, using ultrasound data. Particular hypotheses for each of the three experiments will be introduced in the chapters dedicated to these experiments (Chapters 4-6; see Sections 4.1.2, 5.1.1, 6.1.1).

3. QMUC ULTRASOUND SYSTEM AND GENERAL METHODOLOGY USED IN THIS WORK

> *"...After you've used a term like tongue height – raise the tongue, lower the tongue, move it to the back or front – you begin to feel that that is what your tongue is actually doing. But you are kidding yourself. This is not really what one does in trying to produce a vowel."*
>
> *(Interview with Peter Ladefoged. Fromkin, 1985, p. 5)*

3.1. Introduction

In this chapter, the methodology used in this research is described and discussed. First, ultrasound as a technique for articulatory speech research is presented, with technical details, a brief history of speech studies with ultrasound, and advantages and disadvantages as compared with other techniques. Then the QMUC ultrasound system is described. A multi-channel ultrasound-EPG-acoustics setup is also presented, as it is used in one part of this research. Finally, the methodology used in this research is described, with the information about ultrasound recording, data analysis, and statistical procedures designed in this study for comparison of ultrasound tongue curves.

3.2. Using ultrasound in linguistic research

3.2.1. Ultrasound scanning, visualisation and data analysis

The procedure of ultrasound scanning[4] is based on the ability of ultrasound waves of very high frequencies to pass through the various body tissues and to be reflected back

---

[4] For a more detailed description of the ultrasound scanning process see Stone (1999) and Stone (2005); also, Hewlett & Beck (2006) contains a description of the ultrasound scanning technique based on the QMUC ultrasound system.

to the ultrasound source due to the disparity in density of different tissues. Both static and dynamic scanning work in the same way. Electric current stimulates a piezoelectric crystal, which emits a sound wave of an ultra high frequency. The same crystal then receives the reflected echo. In order to obtain the image of a section rather than a single point, mechanical sector scanners or array transducers are used. In the scanner, multiple crystals rotate about a hub, emitting and receiving ultrasound waves. A mechanical scanner or an array transducer is placed below the chin, enabling the ultrasound waves to go upwards and cover the field of view shaped as a sector of 60 to 180 degrees. Ultrasound waves generated by the crystals in the transducer propagate through the soft body tissues. When the wave reaches an interface between substances with tissues of a different density (like muscle-air or muscle-bone) it is reflected back, detected and processed by computer, and then converted to a grey scale visual display. When the ultrasound wave reaches the air at the tongue surface, it is reflected back. A bright white line corresponding to the air immediately above the tongue surface appears on the display. Muscles appear on the scans as dark areas.

Tongue scanning can be performed in two ways: midsagittal and coronal. Midsagittal scanning is more widely used in speech research, as it produces an image of the tongue surface stretching over several centimetres from root to tip, and allows for measuring the displacement of the functionally important regions of the tongue.

A midsagittal image of the tongue is shown in Figure 3-1. The white line shown by arrows is the result of the air interface at the surface of the tongue, where the ultrasound wave reflects back to the source. The black area immediately below the bright white line is the tongue body. The tip of the tongue is on the right. Shadows created by the bones appear on the scan as black areas (the shadow of the hyoid bone is shown by the dotted line on the left, and the shadow of the chin is shown by the dotted line on the right).

Frequencies that have been used for tongue imaging range from 3 MHz (e.g., Stone et al. 1983) to 7.5 MHz (e.g., Shawker et al. 1984).

Various software has been designed in order to facilitate visualising tongue images on the computer screen and to measure tongue contours. “EdgeTrack” software was designed in the Vocal Tract Visualization Laboratory, Dental School, University of

Maryland, Baltimore (e.g., Li et al. 2003; Li et al. 2005). “EdgeTrack” is an automatic system for extraction and tracking of tongue contours from the ultrasound image. The same laboratory designed and introduced the software for displaying tongue surfaces in the three-dimensional space, called “SURFACES” (e.g., Parthasarathy et al. 2005). “Ultrax edge” tracking analysis software was developed at the Interdisciplinary Speech Research Laboratory, Department of Linguistics, University of British Columbia in Canada (e.g., Gick & Rahemtulla 2004). The researchers from the Graduate Department of Speech-Language Pathology at the University of Toronto developed the software called “Ultra-CATS”, for semi-automated analysis of ultrasound data (e.g., Bressman et al. 2005). The software used with the Queen Margaret University College ultrasound system is described below (Section 3.3).

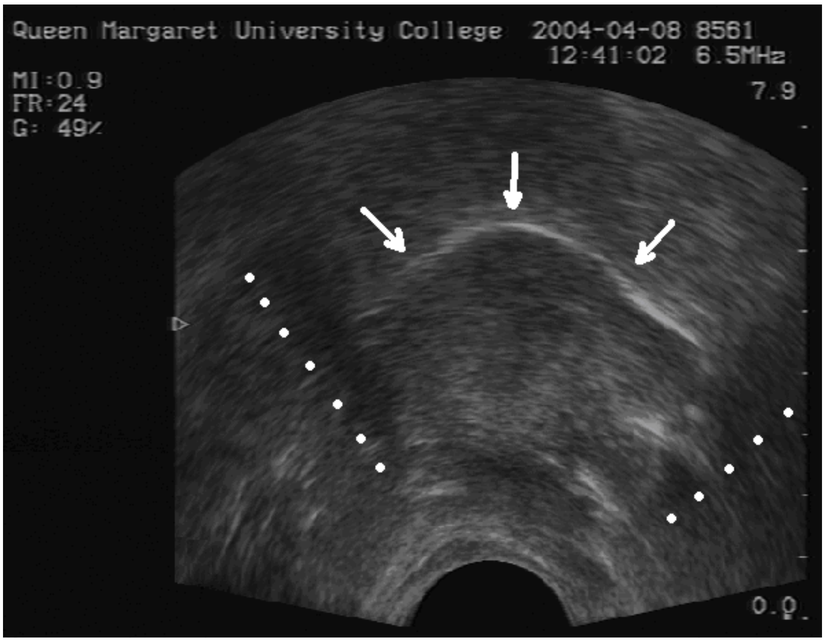


Figure 3-1. Midsagittal scan of the tongue, example of an ultrasound frame. The white line represents air directly above the surface of the tongue. Tongue tip is on the right. Tongue contour is indicated by arrows, the shadow of the hyoid bone is indicated by a dotted line on the left, the shadow of the chin is indicated by a dotted line on the right.

The techniques for measuring and quantitatively comparing tongue curves in the two-dimensional space are generally of two different types: comparing tongue displacement along a measure bar imposed on the tongue curve, and comparing whole tongue contour images. The former procedure is described and applied in, e.g., Vazquez Alvarez et al. (2004), Hewlett et al. (2004), Gick et al. (2006). The latter technique, involving various types of calculations for whole curve comparison, is used in, e.g., Davidson and Stone (2003), Davidson (2005a), Zharkova and Hewlett (2006), Davidson (2006b). The present research employed both the former and the latter techniques, and they will be described below (Sections 3.4.6 and 3.4.7 for comparing whole tongue curves; Section 5.2.8 for measuring tongue displacement along a measure bar).

3.2.2. Application of ultrasound in linguistic studies

Ultrasound imaging is a technique originally used in medicine. It has proved to be useful in studying tongue movements for linguistic purposes, as the technique allows for visualising the internal soft tissues of the articulators involved in speaking. Being non-invasive, harmless and safe, ultrasound imaging can be used extensively, making it possible to get large amounts of data for research. The use of real-time ultrasound allows us to capture the tongue in motion, which can be extremely useful in studying dynamic processes of the language.

The first ultrasound studies of speech date back to the late 60s – mid 70s (Kelsey et al. 1969; Minifie et al. 1971; Skolnick et al. 1975; Zagzebski 1975). For the last three decades, the number of works using ultrasound technology for studying speech has been continuously increasing. Some studies are primarily focused on methodological issues concerned with applying ultrasound to analysing speech: see Niimi and Simada (1980), Sonies et al. (1981), Keller and Ostry (1983), Morrish et al. (1984), Munhall and Ostry (1985), Shawker et al. (1985), Stone et al. (1987), Stone and Lele (1992), Unser and Stone (1992). Some works, presenting linguistically interesting ultrasonic data, relevant for this research, are briefly described below.

A considerable number of works aim to describe tongue position in adults during the production of individual sounds – consonant and vowel phonemes isolated and in a limited context: see MacKay (1977), Stone et al. (1983), Shawker et al. (1984), Morrish et al. (1985), Stone et al. (1988), Stone et al. (1992), Stone (1995), Stone and Lundberg

(1996), Whalen and Gick (2001), Oh (2002), Al-Halees (2003), Wrench and Scobbie (2003), Benus (2005), Scobbie and Sebregts (2005), Brugman (2005), Scobbie and Stuart-Smith (2006).

Some analysis of coarticulation has been done using ultrasound. Parush et al. (1983) was an ultrasound study of lingual coarticulation in vowel-consonant-vowel sequences with velar consonants and back vowels. The researchers used a single pulse-echo ultrasound transducer, it had one beam, and could only track the tongue movements along one line (the beam was directed towards the place of closure for velar stops). Other studies described in this paragraph used a mechanical sector scanner for tongue imaging. Stone (1990), using ultrasound and X-ray data, described tongue movements in American English VCVCə sequences with the consonants /s/ and /l/, and the vowels /i/, /a/ and /o/; Stone proposed a three-dimensional model of tongue movement, based on those data (see more on the model of three-dimensional tongue movement in Stone, 1991). Stone and Vatikiotis-Bateson (1995), using similar data to those described for Stone (1990), aimed at studying how the tongue during the coarticulation process adjusts its position to compensate for conflicting coarticulatory demands. Stone and Vatikiotis-Bateson (1995) used EPG data, additionally to ultrasound data. Tongue behaviour in British English VCV sequences with symmetrical vowels and bilabial consonants was studied with ultrasound and presented in Vazquez Alvarez et al. (2004) and Hewlett et al. (2004). In Wilson (forthcoming), coarticulatory effects of post-velar consonants on long and short vowels /i/, /i:/, /u/, /u:/, /a/ and /a:/ in Nuuchahnulth were analysed with ultrasound, in addition to auditory and acoustic analysis. In Davidson and Stone (2003), Davidson (2005a) and Davidson (2006a), emergent epenthetic schwa was studied in English consonant cluster sequences not permitted by the language phonotactics (for example, /zgomu/, produced as [zəgomu]). In Gick and Wilson (forthcoming), sequences of high tense vowel and liquid were studied, and patterns of an intervening excrescent schwa were observed, using data from English, Beijing Chinese, Nuuchahnulth, Chilcotin (Athapaskan), and Korean (see also Gick and Wilson 2001, Gick et al. 2001).

Ultrasound has also been used for imaging tongue rest position (e.g., Stone et al. 1983), studying articulatory settings across languages (e.g., Gick et al. 2004; Wilson 2005), and suprasegmental influence on segmental structure of speech (Davidson 2005b; Ménard et al. 2005). Ultrasound imaging has been applied to studying second language acquisition (Gick et al. forthcoming), speech errors and speech disorders (e.g., Keller 1987; Bernhardt et al. 2003; Bernhardt et al. 2005; Frisch 2005; Gibbon and Wolters 2005).

3.2.3. Advantages of ultrasound as an articulatory speech research technique

There are a number of advantages in ultrasound technology as applied to speech research. Most importantly, the technique is safe and non-invasive. So unlike, e.g., X-ray imaging, ultrasound allows for recording large amounts of data. This makes it very attractive for using in scientific research.

An advantage of ultrasound as compared with EMA is that the latter is a point-tracking technique. Ultrasound allows for viewing most of the tongue contour (although see Section 3.2.4 for possible problems with imaging the tongue tip). EMA only images the displacement of flesh points. Fowler and Brancazio (2000) describe EMA as "only providing position information about discrete points on the tongue, which may or may not include the tongue portion used to make a consonantal constriction" (Fowler & Brancazio 2000, p. 4). One more side of this problem is that the coil which allows recording of the back part of the tongue is traditionally placed as far as the subject can tolerate, which might happen to be not as far back as the point of interest. Among other disadvantages of EMA are its high cost and invasiveness.

EPG provides only indirect evidence of the tongue behaviour: e.g., the "trough" in the electrode activation pattern during the intervocalic consonant of a VCV sequence is interpreted as corresponding to a relaxation of the tongue position coinciding with the consonant. Ultrasound offers a direct representation of tongue movements in speech.

An advantage of ultrasound is that it can be used together with other methodologies for studying speech. There are possibilities of synchronising the ultrasound signal with acoustics, with video, and also with EPG. In this research, an experiment using combined ultrasound, acoustic and EPG signal has been carried out (see Chapter 6). Potentially, ultrasound can be synchronised with EMA. This combination would provide information on how individual fleshpoints are displaced in relation to the whole tongue contour movement.

3.2.4. Technical challenges of ultrasound as an articulatory speech research technique

As with any other method of studying articulation, ultrasound scanning has its disadvantages. One of the challenges is the difficulty in controlling transducer position relative to the head. This challenge has been dealt with in a number of ways. Special support systems have been designed and applied (e.g., Stone & Davis 1995; Hewlett et al. 2004; Gick et al. 2005). In many studies, a head and transducer support system is used, that helps to keep head and transducer positions fixed, enabling repetitive measurements of the same sound without changing the position of the tongue in relation to the head (for a detailed description of such a system see, Stone and Davis 1995, and Stone 2005). Also, Optotrak has been used in a number of studies. Optotrak is an infrared tracking system that allows for tracking the movements of infrared-emitting diodes. The markers are normally attached to an object that is assumed to be fixed relative to the subject's skull, such as a pair of glasses, and more markers are attached to the ultrasound transducer. The position of the head relative to the transducer can then be controlled for. Another way of keeping the transducer stable relative to the head is when a special chair is used, with a fixed headrest, and a rigid mechanical arm that holds the probe in a fixed position against the subject's neck. Experimental studies have proved that this technology can be successfully used in field work (e.g., Gick et al. 2005; Gick 2005). For the description of the head and transducer support system used in the present work, see Section 3.3.1.

Another challenge of ultrasound is that, unlike, e.g., X-ray or Magnetic Resonance Imaging (MRI), it does not allow us to see the complete picture of all the articulators. No structures can be seen in the image, to which the tongue position could be referred: neither the palate, nor the pharyngeal walls. The shadow of the hyoid bone and the shadow of the chin are present in the picture as black areas. Several ways to display the hard palate have been proposed, among them scanning the tongue pressed against the hard palate, or recording the tongue position during swallowing. The problem of lack of visible fixed reference points within the vocal tract can be partially solved by introducing external points of reference into the tongue video image, like lines, dots, or a grid. Also, the arc where the ultrasound transducer contacts the neck, has been used as a reference point (e.g., Gick 2002; Gick et al. 2006).

The tongue tip and epiglottis cannot be seen on the display either, because of the cavities creating shadows on the image; this should be taken into account when the research is aimed at studying the tongue behaviour in these particular regions. However, it has been proved that the tongue tip may be seen on the screen if the tongue is resting against the floor of the mouth, or if the mouth is filled with saliva (Stone 1999). Use of a wide angle probe is helpful in featuring the tip of the tongue. Also, a better image of the tip can be reached by changing the angle of the transducer as related to the chin (e.g., Gick 2005).

A very good support for ultrasound data, especially in dealing with the problem of reflecting tongue tip activity, is a simultaneous recording of EPG data, synchronised with the ultrasound signal. The patterns of electrode activation on the artificial palate give us information on where precisely the tongue is touching the palate, in addition to the information on the shape of the tongue at that moment (for details about EPG, see, e.g., Hardcastle et al., 1989; Hardcastle et al., 1991a; Hardcastle et al., 1991b; Gibbon et al., 1993; also, see Scobbie et al., 2004, for a description of possible applications of synchronised EPG and ultrasound).

One more challenge that ultrasound presents to the researchers, is its relatively low sampling rate, since the ultrasound image is produced onto the scanner screen with the video output rate. The rate can vary from 25 to 30 frames per second, depending on the country: for example, in most American ultrasound systems, the sampling rate is 30 Hz, and in the UK, in the QMUC ultrasound system, the sampling rate is 25 Hz. This frame rate is obviously quite slow for speech, where precise timing of articulations within a few milliseconds is often linguistically relevant.

There are a few different approaches to the problem of low sampling rate. One of them, taken, e.g., in Davidson (2005a), consists in tracking tongue changes over time, in order to set a context for the sound of interest. In that study, the target sound was an epenthetic schwa. The target vowel in the study had the duration of approximately one ultrasound frame. Davidson wrote: "Because the duration of schwa hovers around 30 ms, it is possible that on any given repetition, the ultrasound machine did not record a frame during the production of the schwa. This makes it risky to assign a specific frame to schwa" (Davidson 2005a, p. 624). In order to avoid this risk, the researcher examined five successive frames, reflecting the change of tongue movement over time corresponding to the target vowel and two surrounding sounds.

Another approach to the problem of the slow sampling rate consists in choosing a sequence of ultrasound frames representing a certain sound (e.g., a liquid consonant), then picking the frame where the gesture is maximally represented, in relation to the other frames, and describing that particular frame, or comparing that frame to the neighbouring frames (e.g., Scobbie & Stuart-Smith 2006).

The approach taken in Vazquez Alvarez et al. (2004) and Hewlett et al. (2004) is based on choosing sounds with relatively fixed steady states lasting for longer than two consecutive ultrasound frames, and then picking a point at the middle of these sounds, taking the corresponding ultrasound frame as being representative of the tongue shape during that sound. This approach is used in the present work, and it will be described in detail in Section 3.4.4 (general description), as well as in Sections 4.2.4, 5.2.4 and 6.2.4 (details for individual experiments).

3.3. QMUC ultrasound system

The ultrasound scans in this work were collected using the Queen Margaret University College ultrasound system, which is described in this section.

3.3.1. Hardware

The hardware consists of a Merlin Ultrasound Scanner Type 1101 and an Endovaginal

End-fire Transducer Type 8561, with a sector of 160 degrees (see *Merlin Ultrasound Scanner Type 1101, User Guide*, for details). The transducer frequency employed in ultrasound studies with the QMUC system is normally of 6.5 MHz. Scanning is performed at a frame rate of 25 frames per second. Head and transducer support system (a helmet with the transducer attached) is used, for immobilising the head in relation to the ultrasound transducer and for capturing ultrasound images with reliability and accuracy. A photograph of the system is presented in Figure 3-2.

Figure 3-2. Queen Margaret University College ultrasound system. The subject is wearing the helmet. The subject and the experimenter are in the sound-treated studio, with the ultrasound scanner and the computer.

3.3.2. Software

The software used with the ultrasound system for data recording and analysis is the programme “Articulate Assistant” (Special ultrasound edition Revision 1.12) designed by the company Articulate Instruments Ltd, and modified for use with ultrasound video images. The programme allows for synchronised recording and playback of ultrasound images and the acoustic signal, together with annotation and analysis routines. For more details of the programme, see Wrench (2006). An example of the ultrasound image as it appears on the computer screen is given in Figure 3-1.

3.3.3. Ultrasound synchronised with EPG

The QMUC ultrasound system is a multi-channel system, in that it has the possibility of recording and analysing synchronised ultrasound, EPG and acoustic signals. EPG data provide information on the patterns of electrode activation on the artificial palate that the subject is wearing, and ultrasound produces an image of the tongue corresponding in time to the EPG signal. The overview of the EPG system is presented in Figure 3-3.

The EPG system used in this work was WinEPG^{TM}, designed by Articulate Instruments Ltd. The hardware is largely based on the EPG3 system, developed at the University of Reading (see, e.g., Hardcastle et al., 1989; Hardcastle et al., 1991a; Hardcastle et al., 1991b; Jones & Hardcastle, 1995). The key feature in WinEPG^{TM} is that it allows for data analysis in Windows, unlike the DOS-based EPG 3 (for details of WinEPG^{TM}, see Scobbie et al., 2004).

Figure 3-3. Overview of the EPG system at QMUC.

During the recording, the subjects wear custom-made artificial palates (Figure 3-3, in front, on the right). There are 62 electrodes embedded in the artificial palate, each electrode individually wired. During the recording, the subject wears the palate and holds in one hand an electrode (Figure 3-3, in front in the middle), which provides electronic circuits. These circuits are scanning the electrodes in the artificial palate during the recording, and when contact of the tongue with an electrode occurs, a signal is conducted to the external processing device, and into the computer. The sampling rate of this EPG system is 100 frames per second.

The computer provides synchronised recording of the EPG signal, acoustics and ultrasound. The software used for simultaneous acquisition and analysis of EPG, ultrasound and acoustic data is a version of the Articulate Assistant programme. An example of an ultrasound image synchronised with EPG and acoustics is presented in Figure 3-4. For details of the software used for combined analysis of ultrasound, EPG and acoustic data, see Wrench (2006).

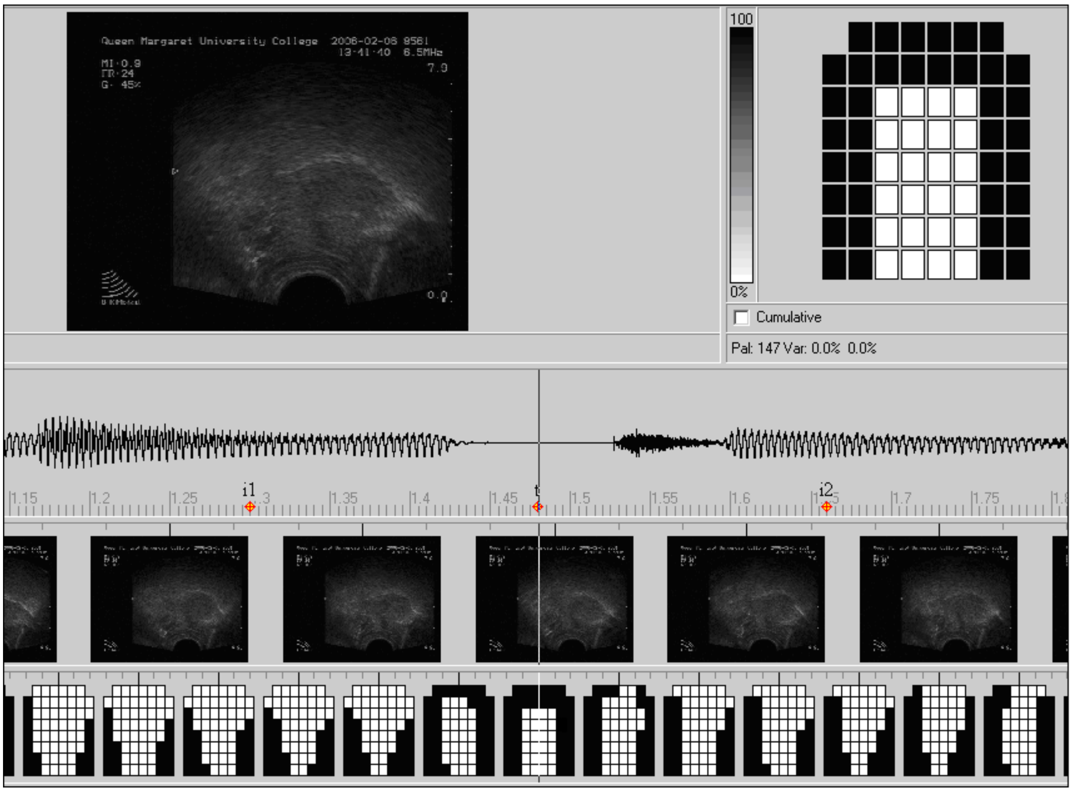


Figure 3-4. Example of an ultrasound frame and an EPG frame synchronised with the acoustic signal. Note that only selected ultrasound and EPG frames are featured below the acoustic waveform. Precise locations of all the ultrasound and EPG frames are indicated by grey ticks above the rows of frames, every 40 ms for ultrasound and every 10 ms for EPG.

The main subject of this work is ultrasound analysis of speech, but ultrasound combined with EPG was used in Experiment 3, and the results are reported in Chapter 6.

3.4. Methodology used in this work

3.4.1. Instrumentation and recording procedure

All the recordings in this study were made in a sound-treated studio in Speech and Hearing Sciences at QMUC. Ultrasound images were collected with the QMUC ultrasound system as described in Section 3.3. The transducer was aimed to be at a right angle to the line of the jaw. Therefore, the probe was expected to be approximately orthogonal to the tongue surface. In Experiments 1 and 2 (Chapters 4 and 5, respectively), the subjects were in the same sound-treated studio as the experimenter. The acoustic signal in Experiments 1 and 2 was recorded using a head mounted dynamic microphone, Shure SM10A (Figure 3-1). In Experiment 3 (Chapter 6), the subjects were in the sound-treated studio, and the experimenter was in the adjacent room. The acoustic signal in Experiment 3 was recorded using an Audiotechnica ATM10a microphone. EPG data were recorded in Experiment 3, simultaneously with ultrasound and acoustic signals, for two speakers.

In all the experiments, the data were VCV sequences embedded in carrier phrases. The subjects read the sentences as they appeared one by one on the computer screen in front of them. The order of presentation was not randomised (see the details of the order of presentation in Experiments 1, 2 and 3 in Sections 4.2.3, 5.2.3 and 6.2.3, respectively). Before the recording, the subjects were given a printout of the sentences, for a quick practice, in order to ensure that they produce the sentences fluently.

Subject recruitment letters, information sheets for participants, and consent forms are presented in Appendix I.

3.4.2. Pilot experiments

Pilot recordings were conducted before each of the three experiments, every time with one subject. This was undertaken in order to test the applicability of the suggested procedure of segmenting the acoustic signal, and to establish explicit acoustic criteria for identifying time points for extraction of ultrasound frames. Pilot experiments showed that the suggested methodology was plausible.

3.4.3. Data analysis: introduction

Below, a description is given of the data analysis procedures that were used in this work, and that are currently used in ultrasound analysis of tongue curves with the QMUC ultrasound system. In the Method section of each experiment, particular details of the analysis procedure, unique to that experiment, are described (see Sections 4.2, 5.2, and 6.2 for Experiment 1, Experiment 2, and Experiment 3, respectively).

3.4.4. Ultrasound software analysis

3.4.4.1 Segmentation and annotation of the acoustic signal

As mentioned in Section 3.4.1, in all three experiments the data were VCV sequences. So there were similar segmentation and annotation criteria throughout this work. These criteria are described below. Specific details of each experiment are presented in the Method section of each relevant chapter.

First, three time points were identified in the acoustic signal (a waveform and a spectrogram): the middle of the first vowel (V1), the middle of the consonant (C), and a particular point in the second vowel (V2). The latter was identified using the following procedure: the interval from mid V1 to mid C was calculated, and then the same length of interval from mid C into V2 was measured. The reasons for choosing a V2 time point in this particular way are as follows. It has been shown in the literature that in VCV sequences, the movement from the consonant towards the following vowel is the greatest and the most rapid during the very early stage of the vowel, and much less so in its later stage (e.g., Fant 1969; Kent & Read 2002). Our ultrasound data also suggest that the most noticeable change from C to V2 is happening soon after the release, and there is not much difference between the ultrasound image at our V2 annotation point and at the acoustic mid-point of V2.

Annotations were placed at the above-named three time points.

3.4.4.2 Spline fitting

The next procedure involved creating three “splines” (curves corresponding to the tongue contour) for each VCV token, at each of the three time points. Splines were created within the ultrasound analysis software, Articulate Assistant, and manually fitted to the tongue curves of interest.

In Figure 3-1, the white line of the midsagttal tongue contour signifies air directly above the surface of the tongue. In this work, splines were fitted to the lower edge of this white line. Graphical illustrations of splines superimposed on the ultrasound image and fitted to tongue curves are presented separately in each experiment description, in Sections 4.2.4, 5.2.4 and 6.2.4.

In Articulate Assistant, splines are defined by a number of “knots”, or points, specified by the user, separately for each spline. Usually, between five and ten knots are enough to describe a curve that fits the contour of the surface of the tongue. Splines created by the user can be of different length. The tongue changes its shape during production of different sounds, and, depending on the tongue position, a longer or a shorter curve may be visible on the screen. This is reflected in spline drawing.

Within Articulate Assistant, splines are stored in pixels. In order to obtain a quantitative representation of the tongue shapes, which could be analysed in Matlab, the splines must be scaled so that measurements relate to the size and distance moved in the real world. It is possible to do this, as the ultrasound video image includes a depth setting scale (see Figure 3-1, on the right). From within the Articulate Assistant software, a line can be drawn on the ultrasound image that matches the length of the scale. Then the length in millimetres of this line can be specified using the software interface. Thus the software “knows” how to scale curves drawn on the ultrasound image to tongue surface curves in the real world measured in millimetres.

Each spline is exported from Articulate Assistant into a text file. Splines are exported as a series of interpolated xy points with the maximum resolution the image allows. The series of xy points is then imported into Matlab for plotting and analysis. An example of the text file with xy data as it is exported from Articulate Assistant is presented in Appendix II (the file contains xy data for the three curves presented in Figure 3-5, Section 3.4.5).

Data points are not equally spaced along the curves. At the moment, it is not possible to control for the spacing of the data points along the curve within the QMUC ultrasound system. There may be occasions when particular regions of the curve are over- or underrepresented by points, and this may affect the average nearest neighbour distance value (see details of the nearest neighbour distance calculations in Section 3.4.6), in that certain parts of the curve are over- or underrepresented in the resulting average distance number. For example, in Figure 3-8 (Section 3.4.6), at the apex of the black curve, the points are spaced less densely than, e.g., further on the right and on the left in this curve. Therefore, the distances between the two curves in the apex region are underrepresented in the nearest neighbour calculations, and in the resulting average distance value. At the moment, this problem in the QMUC ultrasound system is being explored, and the solution is yet to be found. Ideally, data points on the curves should be equally spaced. However, I do not think that the problem described here seriously affects the validity of the analysis, given the great number of data points representing tongue contours. The issue of unequal spacing would have been more important if the number of data points representing each curve had been much smaller than that used in this work.

Below, several key Matlab analysis procedures are described, that were used recurrently in this work. For details of implementing these procedures in particular experiments, see the Method section of the experiments (Sections 5.2 and 6.2).

3.4.5. Plotting tongue curves in Matlab

The tongue curves are presented in this work as Matlab-generated graphs. One Matlab-based step in curve analysis used in all the experiments of this study was plotting tongue curves. The procedure of producing graphs in Matlab is described in this section.

First, the xy data for each curve are imported from the text file into Matlab. The graphs are then plotted. Matlab Scripts were developed for plotting curves in various ways. For example, tongue contours for a single three-spline token can be plotted, as in Figure 3-5.

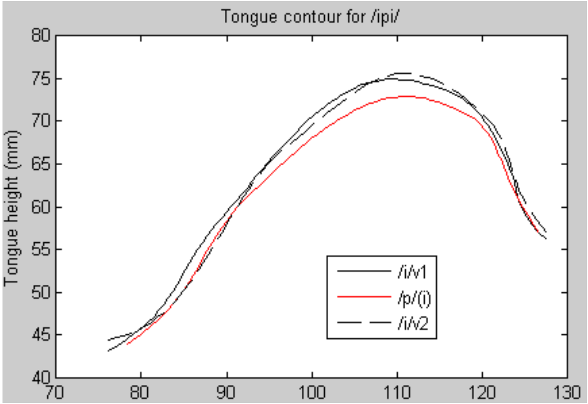


Figure 3-5. Example of three tongue curves taken from a single /ipi/ token, produced by one subject. Solid black line – /i1/, red line – /p/, dashed black line – /i2/.

Also, tongue contours can be plotted for 15 repetitions of the same stimulus. Examples of plots are given in Figure 3-6 (each plot is from a different speaker). The plots represent 15 repetitions of the same sound in the following three conditions: in one context (e.g., 15 repetitions of /k/ in the context of /a/, Figure 3-6a), in two different contexts (e.g., 15 repetitions of /h/ in the context of /i/ and /a/, Figure 3-6b), or in three different contexts (e.g., 15 repetitions of /p/ in three different vowel contexts, Figure 3-6c).

Average tongue curves over 15 repetitions can also be plotted (see Figure 3-7). The method of arriving at an average curve is described below. The 15 curves representing 15 repetitions of the same sound are imported into Matlab as sets of x and y values (as described in Section 3.4.4.2). The first xy pair (from the leftmost end of the curve) is taken from all the 15 curves; 15 x values are averaged, resulting in one number, and 15 y values are also averaged, resulting in one number. These average x and y values become the coordinates of the first data point on the average curve. Then the same calculations are carried out with every consecutive xy pair of the 15 curves, producing a set of xy values representing the average curve.

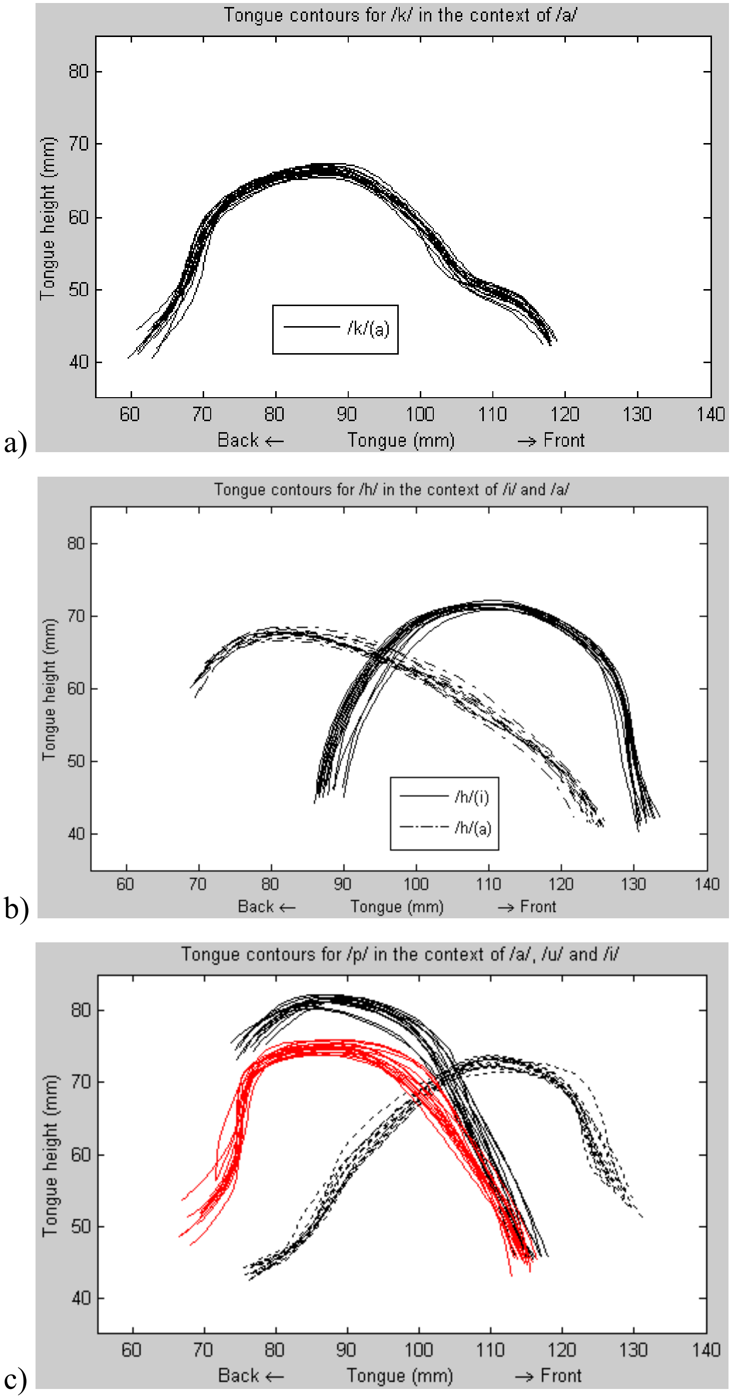


Figure 3-6. Examples of tongue curve plots featuring 15 repetitions of the same stimulus, produced by one subject: a) in one context; b) in two different contexts; c) in three different contexts. Note that the three plots represent productions by three different subjects.

The 15 curves to be averaged vary in length, i.e., in the number of xy data points (see Section 3.4.4.2). The averaging procedure is therefore carried out with xy data points until the last xy point on the shortest curve of the 15. The remaining xy points on the other (longer) curves are not included in the averaging procedure.

It is debatable whether this is an entirely legitimate process for finding an average between several curves, because the curves to be compared are generally of slightly different length, and because the 15 y values averaged do not necessarily correspond to exactly the same point on the x axis. However, this procedure seems to be a reasonable way of portraying an average curve, from a set of rather similar looking curves, representing repetitions of the same stimulus by the same subject. We should also note here that no numerical measurements were carried out on the averaged curves. Average curves were used in this work only for qualitative analysis and illustrative purposes rather than quantitative analysis.

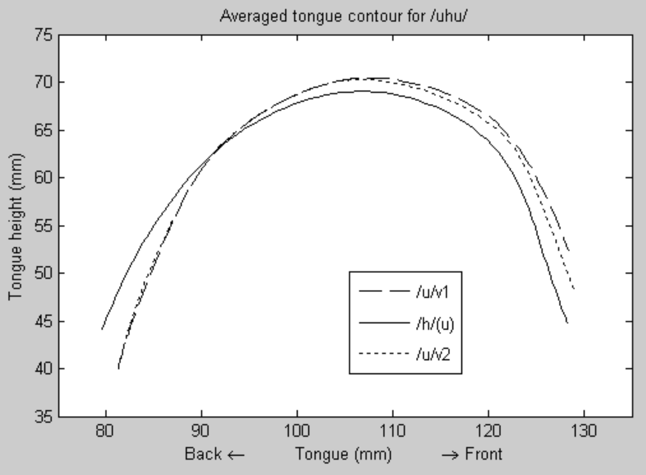


Figure 3-7. Example of three average tongue curves taken from 15 repetitions of /uhu/ produced by one subject.

3.4.6. Comparison of curves: the Nearest Neighbour technique

In order to quantitatively compare different ultrasound tongue curves, the technique called “Nearest Neighbour” was introduced. The technique is illustrated below, using two curves displayed in Figure 3-8 as an example (in actual calculations, more curves are compared with each other; this will be illustrated in Section 3.4.7).

Each point on each curve is compared with each point on the other curve, by means of calculating a distance between the two points.

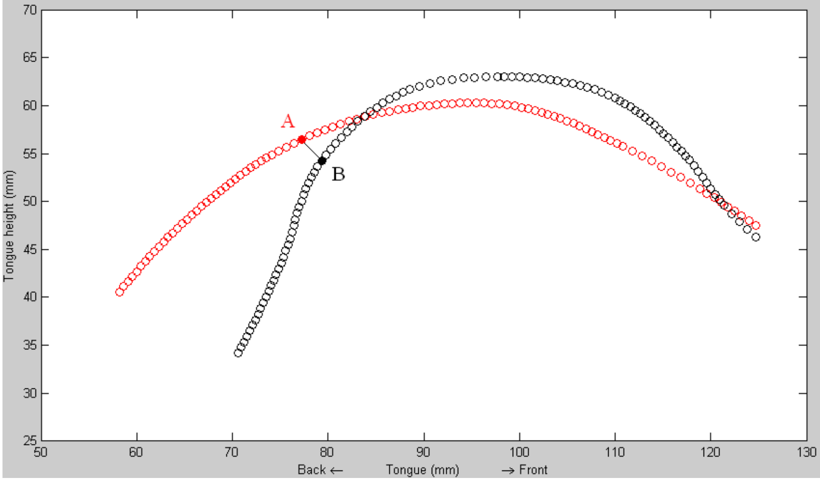


Figure 3-8. An example of two curves plotted in Matlab. Two highlighted points (Point 37 on the red curve, labelled "A", and Point 34 on the black curve, labelled "B") exemplify calculation of the distance from Point A on the red curve to its nearest neighbour on the black curve.

In Figure 3-9, the distance calculation is illustrated. The two points, A and B, are taken from the two curves in Figure 3-8. Each point has its x and y values. The distance between the points is equal to the hypotenuse of the triangle (h), the two legs of which are represented by the lines j and k. The formula used for obtaining the distance value is the following: $\boldsymbol{h} = \sqrt{(\boldsymbol{j}^2 + \boldsymbol{k}^2)}$.

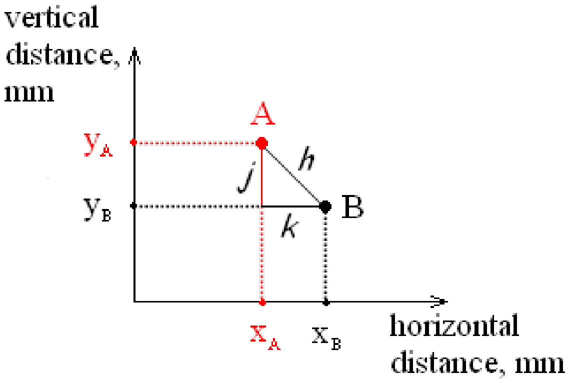


Figure 3-9. Illustration of the calculation of the distance between two points belonging to different tongue curves (an enlarged section of Figure 3-8).

When all the distances are calculated from the first point on Curve 1 to each point on Curve 2, the shortest of these distances is called the nearest neighbour distance. Then the procedure is repeated for the second point on Curve 1, then the third point on Curve 1, and so on until the last point on Curve 1. After that, the nearest neighbour distances are calculated from the first point on Curve 2 to each point on Curve 1, from the second point on Curve 2 to each point on Curve 1, and so on, until the last point on Curve 2. This is done because the curves can be of different length (as noted in Section 3.4.4.2); by comparing Curve 1 to Curve 2 *and* Curve 2 to Curve 1, we obtain slightly differing lists of nearest neighbour distance values. In the end, all these values are averaged, giving us one number, representing the distance between the two curves.

3.4.7. Statistical comparison of curve sets

The Nearest Neighbour technique, described in detail in the previous section, was repeatedly used in this work for comparing sets of ultrasound tongue curves. One procedure, calculating average distance between two sets of tongue curves, is described in this section.

The aim is to measure the distance between two sets of tongue curves (15 repetitions in each set), taken from two particular time points, e.g., the middle of the /t/ closure in two different vowel contexts, as in Figure 3-10.

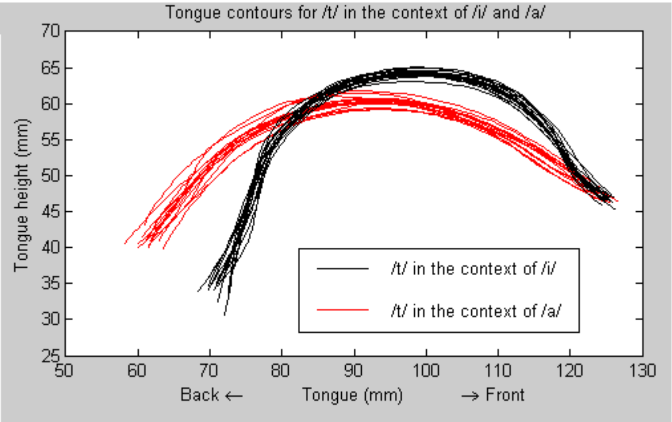


Figure 3-10. Example of two sets of curves (15 repetitions in each set), which are compared in Matlab.

3.4.7.1 Calculating a distance between two curve sets

First, the xy data for the relevant sets of tongue curves are exported from the ultrasound software into a text file, and then imported into Matlab. Next, a list of average nearest neighbour distance values is calculated, using the protocol called “across group comparison”, described below and illustrated in Figure 3-11.

ACROSS GROUP COMPARISON

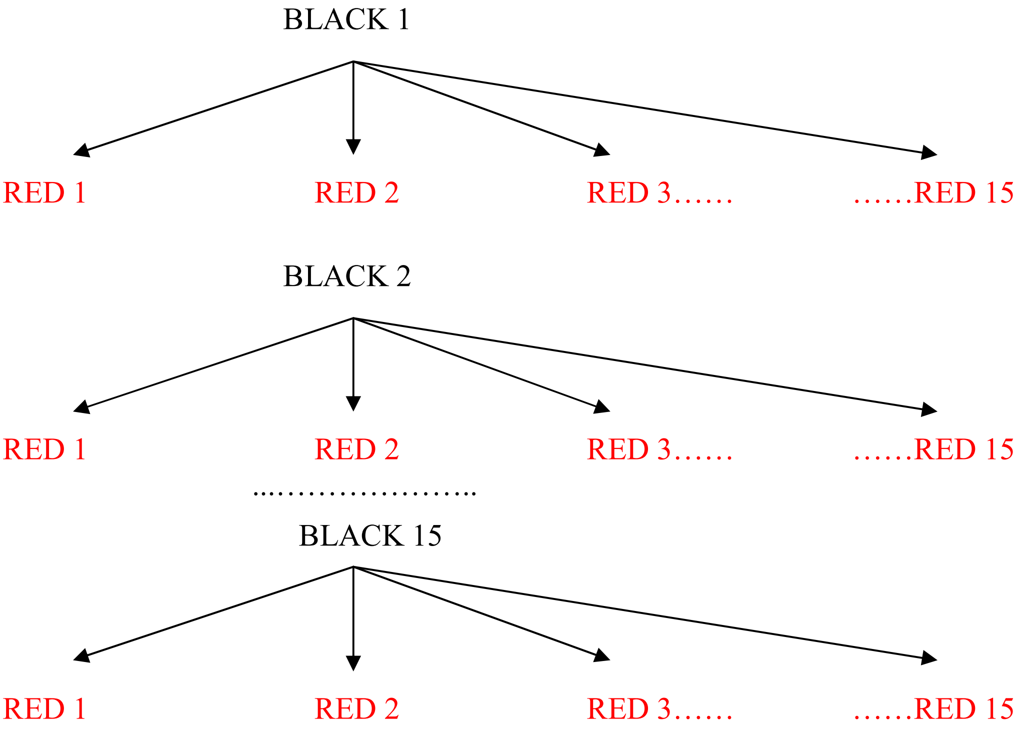


Figure 3-11. The procedure of “across group” comparison of one set of curves with another.

A representation of the curves that are compared with each other in “across group” calculations is given in Table 3-1, as a matrix. An example of the matrix with average nearest neighbour distances between two actual sets of curves (plotted in Figure 3-10) is presented in Appendix III.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
2	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
3	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
4	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
5	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
6	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
7	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
8	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
9	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
10	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
11	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
12	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
13	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
14	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
15	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

Table 3-1. A matrix with pairs of curves for “across group” comparison. The numbers in the first column represent the 15 curves in the black set, and the numbers in the first row represent the 15 curves in the red set. In total, 225 average nearest neighbour distance values are obtained. Each cross represents the absolute difference between the two curves concerned, using Nearest Neighbour.

Using the Nearest Neighbour technique (described in Section 3.4.6), each of the 15 black curves is compared with all the red curves. The resulting list of average nearest neighbour distances consists of 225 values (M*N, where M is the number of curves in the first group and N is the number of curves in the second group). Then these values can be averaged, to produce a single number representing the average distance between the two sets of curves, and a standard deviation may be obtained.

3.4.7.2 Calculating the significance level of a difference between two curve sets

The procedure of comparing two sets of curves for significant differences is described in this section. The aim is to compare for significant differences two sets of tongue curves (15 repetitions of each), taken from two different contexts, e.g., the middle of the /t/ closure in two different vowel contexts, as in Figure 3-10.

First, the xy data for the relevant sets of tongue curves are exported from the ultrasound software into a text file, and then imported into Matlab. Next, three lists of average nearest neighbour distance values are calculated. One list is created using the "across group comparison" protocol described in the previous section. A further two lists are made using the protocol called "within group comparison", described below.

The lists of "within group" nearest neighbour distances are calculated separately for the set of black curves and for the set of red curves, in the way illustrated in Figure 3-12. Using the Nearest Neighbour technique (see the description in Section 3.4.6), each of the 15 black curves is compared with all the other black curves. The resulting set of average nearest neighbour distances consists of 105 values. The number of values is N*(N-1)/2, where N is the number of curves in the set. The same procedure is carried out with the set of red curves, resulting in another set of 105 values.

WITHIN GROUP COMPARISON

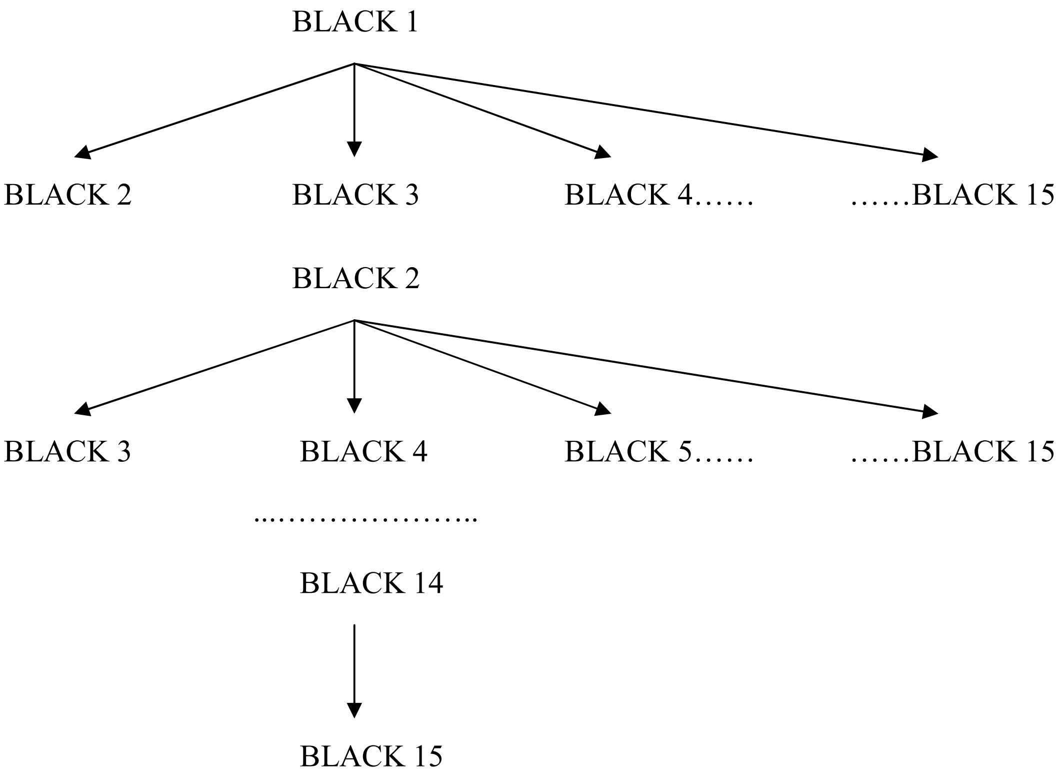


Figure 3-12. The procedure of "within group" comparison of 15 curves.

A representation of the curves that are compared with each other in "within group" calculations is given in Table 3-2, as a matrix. Examples of matrices with average nearest neighbour distances within each of the two sets of curves (plotted in Figure 3-10) are presented in Appendix III.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1		x	x	x	x	x	x	x	x	x	x	x	x	x	x
2			x	x	x	x	x	x	x	x	x	x	x	x	x
3				x	x	x	x	x	x	x	x	x	x	x	x
4					x	x	x	x	x	x	x	x	x	x	x
5						x	x	x	x	x	x	x	x	x	x
6							x	x	x	x	x	x	x	x	x
7								x	x	x	x	x	x	x	x
8									x	x	x	x	x	x	x
9										x	x	x	x	x	x
10											x	x	x	x	x
11												x	x	x	x
12													x	x	x
13														x	x
14															x
15															

Table 3-2. A matrix with pairs of curves for "within group" comparison. The numbers (1 to 15) both in the first row and in the first column represent 15 repetitions of the same stimulus, and the matrix illustrates which pairs of curves are compared. In total, 105 average nearest neighbour distance values are obtained. Each cross represents the absolute difference between the two curves concerned, using Nearest Neighbour.

The three lists of average nearest neighbour distance values (one across-group and two within-group) are compared for significant differences, by conducting a one-way Univariate ANOVA in SPSS, for comparing within-group variation for the set of black curves, within-group variation for the set of red curves, and across-group variation. If a Post Hoc test shows a significant difference between across-group variation and both within-group variations, at the 0.05 level, then the distance between the two sets of curves is considered significant.

3.5. Advantages and challenges of using ultrasound in this research

Specific advantages of using ultrasound for studying lingual coarticulation in VCV sequences are that, being safe, ultrasound allows for collecting as many data samples as needed, to account for the variation in the production of the sequences and to conduct reliable statistical analyses. A particular advantage of using ultrasound in this project is that it produces an image of almost the whole contour of the tongue. This can be crucial when the location of the part of the tongue "specified" for producing a particular movement is not obvious. In X-ray imaging, the whole tongue contour image is also produced, but the nature of the X-ray scanning imposes limitations on its use, due to safety requirements. Magnetic Resonance Imaging (MRI) produces the whole tongue contour, but this technique does not have the temporal resolution allowing for studying dynamic speech events. EMA only allows us to see traces of the few coils placed on particular points of the tongue, and these points, as noted in Section 3.2.3, may not be exactly where a certain tongue movement occurs. EPG is an indirect measure of tongue movements, as it reflects contact patterns of the tongue with the hard palate.

One of the challenges of ultrasound tongue imaging, described above (Section 3.2.4), is the difficulty in controlling for the transducer position relative to the head. The head and transducer support system used at QMUC has proved to provide stable data throughout the recording session. It has also been shown that two separate recording sessions with the same data give similar results (Hewlett et al. 2004).

Another challenge, that is rather serious in coarticulation research, is the relatively low frame rate of ultrasound (25 frames per second). In the data analysis, there is a new ultrasound image every 40 ms, which is a considerable amount of time in speech production. So there is no information on the precise timing of the changes that occur in the tongue shape during the 40 ms of each frame. This can be a problem for studying fast tongue transitions, as in diphthongs, when it is important to trace the changes happening over the time that may correspond to only two or three ultrasound frames. The slow frame rate does not create difficulties for analysing VCV coarticulation using the criteria employed in this work, because each of the three segments of interest, V1, C and V2, lasts on average over 80 ms, and each of them is represented by one annotation point, corresponding to one ultrasound frame.

The timing challenge described in the previous paragraph can be tackled using the combined methodology of ultrasound and EPG. The frame rate of EPG is 100 frames per second. Analysing synchronised data from ultrasound and EPG has two big advantages.

First, it provides information on the tongue shape *and* the precise location of the tongue contact with the hard palate. Secondly, four times more information on the fine detail of the timing of articulations is obtained than with ultrasound alone. The methodology of using these two techniques combined with acoustics was tested in this work, and is described in Chapter 6.

3.6. Summary

In this chapter, ultrasound as an experimental technique for studying speech was described. Principles of ultrasound scanning of tongue movement were outlined, a short background to using ultrasound in speech research was presented, and advantages and disadvantages of this technique were given.

The QMUC ultrasound system was described, and the general methodology used in this work was outlined. References to this chapter will be found throughout the thesis, as it describes three important constituents of the study: the ultrasound system hardware, the software used for the first stage of the data analysis, and the key Matlab-based procedures used repeatedly in this work at the second stage of the data analysis. These procedures include the Nearest Neighbour technique and its application for implementing two analysis procedures: calculating distances between two sets of tongue curves, and comparing two sets of tongue curves for significant differences. These procedures are currently used with the QMUC ultrasound system for quantitative comparison of different tongue configurations in speech, and their description may be useful for potential users of the ultrasound system.

4. EXPERIMENT ONE:
RUSSIAN BILABIAL STOPS IN VCV SEQUENCES

4.1. Introduction

This experiment addressed lingual coarticulation in Russian VCV sequences with labial consonants. The notion of Coarticulation Resistance (CR) is used in interpreting the results.

It is known from previous research (e.g., Recasens 1999) that bilabial consonants have a very small degree of resistance to lingual coarticulation, i.e., they are strongly coarticulated with neighbouring sounds, as the tongue does not need to take any particular shape for their production.

Even though a large degree of lingual coarticulation in labial consonants has been demonstrated cross-linguistically, there are some contradictory data in the literature. A finding that has been consistently reported in the literature for many years, is that in VCV sequences, there occurs a “discontinuity in coarticulation” (Perkell 1986), or a change in tongue position, between the vowels (see Section 2.2 for details). However, the discontinuity in coarticulation in VCVs with bilabial consonants has been reported to occur in some languages, but not in others (for the description, see Section 2.2.12).

This chapter reports the results of an investigation of lingual coarticulation in VCVs with labial consonants in Russian, a language where it has not been studied before. Large coarticulation of the consonants with the surrounding vowels is expected, based on the existing cross-linguistic evidence (e.g., Recasens 1999). Predictions concerning fine articulatory details of tongue behaviour during a VCV sequence will be made in Sections 4.1.1 and 4.1.2. As we will see below, evidence from the literature suggests that particular coarticulatory patterns in Russian may differ from the languages where VCVs with bilabial consonants have been studied.

4.1.1. Relevant characteristics of Russian consonants and vowels

It is interesting to investigate coarticulation in VCV sequences in Russian, because of some differences in the phonological system between Russian and the languages where lingual coarticulation in VCV sequences with bilabial consonants has been studied so far. This section describes some phonological and phonetic characteristics of Russian, which are relevant for the present study.

The principal difference between Russian and the languages described in Section 2.2.12, is that Russian has a phonological opposition in palatalisation (e.g., Zinder 1979; Bondarko 1998; Kasatkin 2006; Reformatskij 2006). Most consonants in the system are organised in pairs, i.e., there are two consonant phonemes that differ only in one feature (for example, /p/ and /p^j/, /b/ and /b^j/, /t/ and /t^j/, /r/ and /r^j/). One of the consonants from such a pair is called “non-palatalised”, or “hard”; the other consonant from the pair is called “palatalised”, or “soft”. From this point these two types will be referred to as palatalised and non-palatalised respectively.

Palatalised consonants have been described in the literature as characterised by a secondary articulation, in addition to the principal articulation. For producing palatalised consonants, the tongue dorsum raises to the hard palate, towards a position that is very close to the lingual position for the vowel /i/ (e.g., Matusevich 1948; Fant 1960; Zinder et al. 1964; Bondarko 1998; Bondarko 2005; Kasatkin 2006; Reformatskij 2006). The secondary articulation is coordinated with the principal articulation; so in palatalised consonants of a different place and manner of articulation, position of tongue dorsum differs, depending on the principal articulation (Zinder 1979; Bondarko 2005).

Non-palatalised consonants have been described as phonetically velarised (e.g., Trubetzkoy 1939; Bondarko et al. 1968; Reformatskij 1970; Zinder 1979; Hamann 2004; Bondarko 2005; Reformatskij 2006). In Fant (1960), Russian non-palatalised consonants were demonstrated to be both velarised and pharyngealised, i.e., the back part of the tongue was displaced upwards and backwards, in addition to the principal articulation; exact position of the tongue dorsum and root varied, depending on the principal articulation. In a recent book on Russian phonetics, Russian bilabial non-palatalised consonants are described as both velarised and pharyngealised (Kasatkin 2006).

VCV sequences that have been studied in other languages, include three different vowel types: /i/, /u/ and /a/ (Section 2.2). This work aims to use similar vowels in Russian VCVs, in order to facilitate cross-linguistic comparison. The peripheral vowels /i/, /u/ and /a/ are used. Stressed allophones of these phonemes produced outwith context are [i], [u] and [a], respectively (e.g., Scherba 1912; Zinder 1979; Kuznetsov 2000).

In Russian, the consonant in the iCi sequence is phonologically palatalised. There are two different theoretical approaches to the question of the phonological status of [i] preceded by palatalised consonants in Russian (for a description of both, see, e.g., Kodzasov & Krivnova 2001). One approach claims that palatalised consonants precede the allophone [i] of the phoneme /i/, while after non-palatalised consonants, there occurs the allophone [ɨ] of the same vowel phoneme. Another approach treats [i] and [ɨ] as separate phonemes, and claims that palatalised consonants only occur before /i/, while the phoneme /ɨ/ can only be preceded by non-palatalised consonants. However, both approaches agree that the consonants preceding [i] are always palatalised. So Russian /iC^{j}i/ sequences are used in this experiment.

Before the vowels /a/ and /u/, both non-palatalised and palatalised consonants can occur. In this work, /aCa/ and /uCu/ sequences with non-palatalised consonants are studied, in order to have more comparable phonetic characteristics with VCVs from other languages.

Phonetic characteristics of the Russian consonants described above have implications for studying tongue behaviour in VCV sequences with bilabial consonants. Given that the tongue dorsum is raised for the production of both non-palatalised and palatalised consonants, we have reasons to expect no discontinuity in tongue behaviour between the two high vowels in VCVs with labial consonants in Russian (e.g., /ubu/ or /ip^{j}i/). In VCVs with low vowels and bilabial consonants (e.g., /aba/), it would make sense to expect intervocalic tongue raising. This experiment is designed in order to test these claims.

Experiment 1: Russian bilabial stops

A phonetic characteristic of the opposition of voiced and voiceless consonants in Russian is that voiced consonants have a negative VOT, i.e., they are phonetically voiced (e.g., Bondarko 1998; Kasatkin 2006). Voiceless consonants have a VOT around zero, i.e., they are unaspirated. In previous studies of VCV coarticulation in other languages, some differences between voiced and voiceless bilabial consonants have been observed. For example, in voiced consonants surrounded by low vowels (in American English), there occur bigger troughs than in voiceless consonants (Section 2.2.7). In voiceless consonants surrounded by high vowels (in German), there occur bigger troughs than in voiced consonants (Section 2.2.8). In Russian, unlike English and German, bilabial consonants have a certain constraint on lingual position for their production, as described earlier in this section. So we could expect this phonologically based constraint to affect the production of both voiced and voiceless consonants, resulting in similar lingual coarticulatory patterns in these two types of bilabial stops in Russian.

Stress in Russian may fall on any syllable of a word. It may also change its position in a word from one morpheme to another. Considering this, in the present study it is important to mark the stress in the data presented to the subjects, in order to avoid the situation where they would randomly choose stress position.

An important phonetic characteristic of Russian is that vowels undergo great changes in duration and quality, depending on their position in relation to stress (e.g., Bondarko 1998; Kasatkin 2006; Reformatskij 2006). Vowel duration and quality is the principal phonetic correlate of stress in Russian, with F_0 and intensity being only secondary cues (e.g., Bondarko et al. 1966). A description of the vowel quantitative and qualitative reduction based on auditory and acoustic analysis is presented in, e.g., Kuznetsov (1997), Bondarko (1998), Kuznetsov (2000). The duration of the vowel immediately neighbouring the stressed vowel is approximately two thirds of the duration of the stressed vowel. The reduction in duration affects the quality of the unstressed vowel. Unstressed allophones of the vowels /i/ and /u/ are described as slightly more centralised vowels, transcribed as [ɪ] and [ʊ], respectively. Unstressed allophones of the vowel /a/ are more variable than unstressed allophones of the two high vowels. There is only one low vowel in Russian, and its allophones have a lot of space where they can vary in quality, without risk of being confounded with other vowels (the nearest two vowels in the Russian vowel space are /e/ and /o/). Substantial qualitative variation of the vowel /a/ depending on the context and on the position in relation to stress has been shown in experimental acoustic and perceptual studies (e.g., Kuznetsov 1997; Kuznetsov 2000). For example, the pre-stressed allophone of /a/ following the consonant /t/ is described as a front, [ɛ]-like vowel.

Based on the literature described above, qualitative reduction of unstressed vowels is expected in this work. In /iC^{j}i/ and /uCu/ sequences, the unstressed vowel should be more centralised than the stressed vowel. In /aCa/ sequences, the unstressed vowel, immediately following the consonant /t/ (see Section 4.2.1, for the experimental data), should be further forward than the stressed vowel.

Another feature of Russian, important when analysing coarticulation in VCV sequences, is syllable structure. It has been demonstrated that in Russian, coarticulation in CV sequences is stronger than in VC sequences. Bondarko (1969) studied acoustic characteristics of CV and VC syllables, and found that in CV syllables, there were greater and longer formant transitions than in VC syllables, making the consonant and the vowel of the CV syllable more “contrasted” than the vowel and the consonant of the VC syllable, and thus making it easier to identify the distinctive features of both the consonant and the vowel from the CV-complex than from the VC-complex. It has also been shown that there is strong coarticulation in CV sequences with a word boundary between the consonant and the vowel (e.g., Zinder et al. 1968). Also, studies of anticipatory coarticulation in Russian have shown that an intervocalic string of consonants is more strongly coarticulated with the following vowel than with the preceding vowel (e.g., Kozhevnikov & Chistovich 1965). The implication of all these findings is that lingual coarticulation in VCV sequences should be influenced by the syllable structure of the VCV, i.e., that there should be some evidence in the data that the consonant is influenced by V2 more than by V1.

4.1.2. Hypotheses

The experiment was designed to test the following hypotheses:

1. **There will be a clearly visible difference between tongue contours for bilabial stops in different vowel environments.** If the hypothesis is supported it will be concluded that Russian bilabial consonants are coarticulated with neighbouring vowels, and that their resistance to lingual coarticulation is lower than maximal. The hypothesis will be refuted if there is no clearly visible difference between tongue contours for bilabial stops in different vowel contexts. If this result occurs, it will be concluded that Russian bilabial consonants are not coarticulated with neighbouring vowels in VCV sequences, and that their resistance to lingual coarticulation is absolute.

2. **There will be a continuous tongue movement between the two vowels in /iC^{j}i/ and /uCu/ sequences: specifically, the highest point of the tongue in the C curve will lie between the two vowel curves.** If the hypothesis is supported it will be concluded that there is no discontinuity in coarticulation in Russian /iC^{j}i/ and /uCu/ sequences. The hypothesis will be refuted if the highest point in the C curve is consistently higher or lower than the two vowel curves. If this result occurs, it will be concluded that there occurs a discontinuity in coarticulation in Russian /iC^{j}i/ and /uCu/ sequences.

3. **The tongue dorsum will raise between the two vowels in /aCa/ sequences: specifically, the highest point of the tongue in the C curve will be above the two vowel curves.** If the hypothesis is supported it will be concluded that there is a discontinuity in coarticulation in Russian /aCa/ sequences, and that the discontinuity is in the predicted direction, i.e., tongue dorsum raising. The hypothesis will be refuted if the highest point in the C curve is consistently between or below the two vowel curves. These outcomes would mean, respectively, that there is no discontinuity in coarticulation in Russian /aCa/ sequences, or that there occurs a discontinuity in a non-predicted direction, i.e., tongue lowering.

4. **There will be similar lingual coarticulatory patterns in voiced and voiceless bilabial stops in Russian VCV sequences.** If the hypothesis is supported it will be concluded that the tongue does not contribute to creating a distinction between voiced and voiceless consonants in Russian. The hypothesis will be refuted if any consistent differences in lingual coarticulatory patterns are observed between VCV sequences with voiced and voiceless stops. This would mean that the tongue contributes to creating a distinction between voiced and voiceless consonants in Russian.

5. **There will be a clearly visible difference between the V1 curve and the V2 curve in /iˈCʲi/ and /uˈCu/ sequences: specifically, the tongue dorsum will be lower in V1 than in V2.** If the hypothesis is supported it will be concluded that the unstressed vowel in Russian /iˈCʲi/ and /uˈCu/ sequences is qualitatively reduced. This hypothesis will be refuted if the tongue dorsum is not visibly lower in V1 than in V2. This result would mean that there is no qualitative reduction of unstressed vowels in Russian /iˈCʲi/ and /uˈCu/ sequences.

6. **There will be a clearly visible difference between the V1 curve and the V2 curve in /aˈCa/ sequences: specifically, the tongue blade and the tongue root will be further forward in V1 than in V2.** If the hypothesis is supported it will be concluded that the unstressed vowel in Russian /aˈCa/ sequences is qualitatively reduced. This hypothesis will be refuted if the tongue blade and the tongue root are not visibly further forward in V1 than in V2. This result would mean that there is no qualitative reduction of the unstressed /a/ in Russian /aˈCa/ sequences.

7. **In a V#CV sequence, the C curve will be more similar to the V2 curve than to the V1 curve.** If this hypothesis is supported it will be concluded that the consonant is more influenced by V2, the vowel belonging to the same syllable of the V#CV, than by V1.

The hypothesis will be refuted if the C curve is more similar to the V1 curve than to the V2 curve, or if the C curve is not more similar to any of the two vowel curves. These outcomes would mean, respectively, that the consonant is more influenced by V1 than by V2, or that the lingual position in the intervocalic bilabial consonant is not influenced by the syllable structure of V#CV sequences.

4.2. Method

4.2.1. Experimental items

The data were the symmetrical nonsense VCV sequences /aba/, /apa/, /ubu/, /upu/, /ibʲi/, /ipʲi/, in the carrier phrases, with /t/ being the left and right context for /aCa/ and /uCu/ sequences, and /tʲ/ being the left and right context for /iCʲi/ sequences. The six following sentences were presented to the subjects. The original Russian version is given first, the phonological transcription and the English translation is presented below each sentence. The target VCV sequences are underlined in the transcription. The subjects were presented with the sentences in Russian orthography. The stressed syllable was in capital letters.

Когда она скажет аПА, ты повторяй за ней.
/kaˈgda aˈna ˈskaʒɨt aˈpa tɨ paftaˈrʲaj za nʲej/
When she says aPA, you repeat after her.

Если она скажет аБА, то Алла Отару скажет про бандану.
/ˈjesʲlʲi aˈna ˈskaʒɨt aˈba to ˈalla aˈtaru ˈskaʒɨt pra banˈdanu/
If she says aBA, then Alla will tell Otar about bandanna.

Если она скажет уПУ, ты тогда про измерения ладоней скажи.
/ˈjesʲlʲi aˈna ˈskaʒɨt uˈpu tɨ taˈgda pra izmʲiˈrʲenʲii laˈdonʲij skaˈʒɨ/
If she says uPU, you mention measuring hand palms.

Когда Ада произнесет уБУ, ты её не осмеивай, ладно?
/kaˈgda ˈada praizʲnʲiˈsʲot uˈbu tɨ iˈjo nʲi aˈsmʲeivaj ˈladna/
When Ada utters uBU, do not laugh at her, all right?

Тогда Ира должна сказать иПИ семь тысяч раз, как задаток.
/taˈgda ˈira dalˈʒna skaˈzatʲ iˈpʲi sʲem ˈtɨsʲit͡ʃ ras kag zaˈdatak/
Then Ira has to say iPI seven thousand times, as a deposit.

Тогда Ирина должна сказать иБИ семь раз, - сказала Ада.
/taˈgda iˈrʲina dalˈʒna skaˈzatʲ iˈbʲi sʲem ras skaˈzala ada/
Then Irina has to say iBI seven times, - said Ada.

4.2.2. Subjects

The subjects were three female native speakers of Russian, between 25 and 35 years old. All the subjects were born and raised in urban areas in the European part of Russia, and spoke Modern Standard Russian, with no outstanding dialectal features. All the subjects had graduate-level experience in phonetics. The author was one of the subjects (namely, S3).

4.2.3. Instrumentation and recording procedure

The details of the recording procedure for this experiment are described in Section 3.4.1, together with the general description of the other two experiments' recording setup.

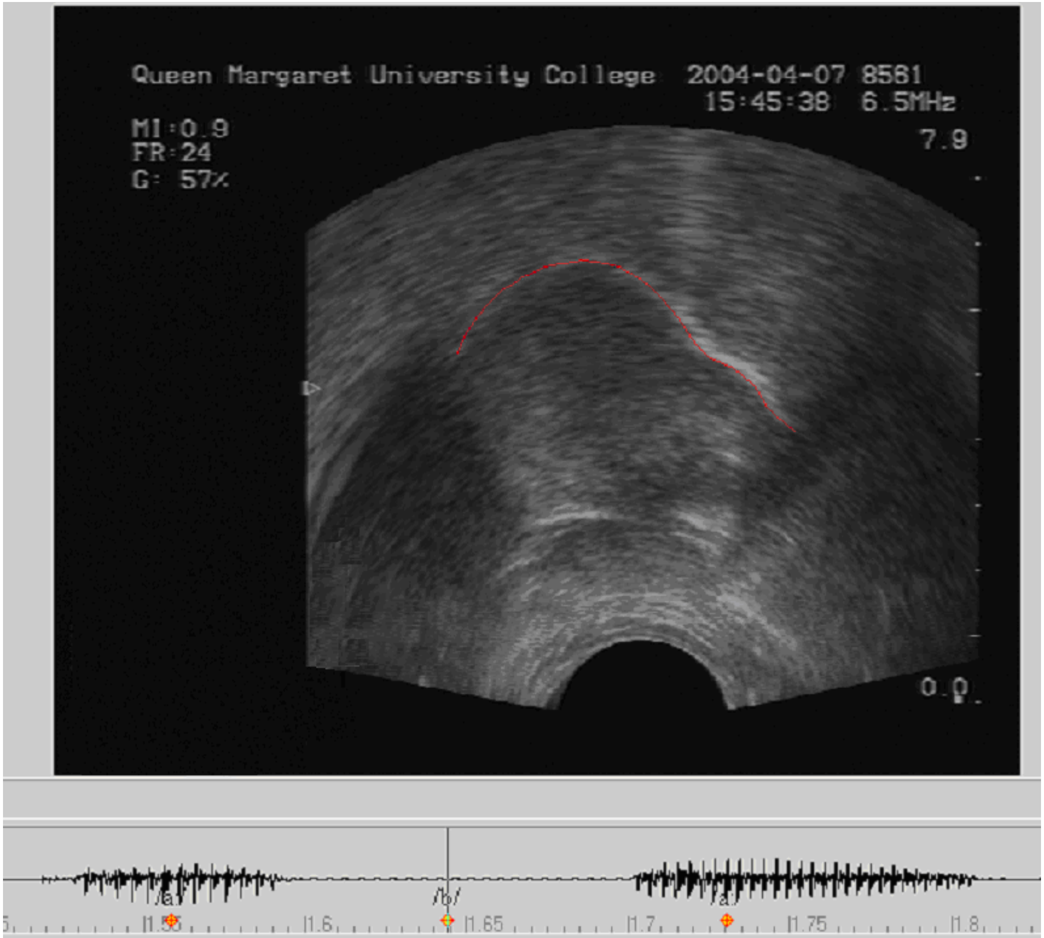
There were 15 tokens in each stimulus type. The total number of VCV sequences recorded and analysed in this experiment was 270 tokens.

The order of presentation was the following. One repetition of each of the six sentences was collected as a block, before moving on to the second block, and so on. The sentences in each block were presented in the same order. The order was as presented in Section 4.2.1.

The participants were given a printout of the sentences, for some pre-recording practice, as described in Section 3.4.1 for all three experiments. The subjects were instructed to produce the capitalised syllables as stressed. They were asked to produce the sentences at a comfortable speaking rate.

4.2.4. Ultrasound software analysis, annotations and splines

Annotating the waveform and creating splines within the ultrasound analysis software (Articulate Assistant), common to all three experiments in this work, was described in Section 3.4.4. In this experiment, the V1 spline was placed at the mid-point of V1, the C spline was placed at the mid-point of the stop consonant closure, and the V2 spline was created at the V2 annotation point described in Section 3.4.4. An illustration of the annotations and spline drawing in this Experiment is given in Figure 4-1.



mid V1 ← mid /b/ → same distance as mid V1-/b/

Figure 4-1. Illustration of the three annotation points and a spline drawn over the tongue contour, corresponding to the /b/ annotation point (the mid-point of the closure).

After annotating the waveform and drawing the splines, the xy data for each curve were exported from Articulate Assistant into a text file, and then imported into Matlab for plotting (see Section 3.4.4.2 on the details of importing tongue curves into Matlab). The Matlab procedures for plotting tongue curves are described in Section 3.4.5.

4.2.5. Qualitative observation of whole tongue contours

In this experiment, visual observation of tongue contours was used as one of the analysis methods. Two different techniques of portraying tongue curves were used. One of them consisted in plotting on the same graph average tongue curves for V1, C and V2 belonging to the same VCV type. Each of the three curves was obtained by averaging 15 curves (see Section 3.4.5), from the 15 tokens of the same sound from a VCV. Another way of displaying tongue curves involved plotting 15 tokens of the same sound on the same graph (see Section 3.4.5). This was done in order to display variation across tokens. In this experiment, this type of plot was used for presenting three sets of 15 tokens of the same sound on the same graph. In one case, three sets of 15 tokens of the same consonant were displayed in three different vowel environments. In the other case, three sets of 15 tokens of the same consonant produced by all three subjects were displayed.

4.2.6. Quantitative distribution of tongue contour sequence patterns

Each VCV token was represented by three tongue curves (V1, C and V2). The criteria used for identifying different relationships among these tongue contours are described in this section. In this experiment, visual observation was used for identifying different tongue contour sequence patterns. Quantitative methods were only used for analysing the distribution of the observed patterns.

The point on the C curve having the maximum y value (or the highest point in the C curve) was taken as a reference point. Then, at the same x value, the y values of the V1 curve and the V2 curve were identified. The location of these two y values in relation to the maximum y value on the C curve was used as a criterion for identifying the tongue contour sequence pattern, formed by the three curves.

There are four theoretical possibilities. In one case, the maximum y value on the C curve is lower than the y values with the same x value on the V1 and V2 curves (informally: the highest point on the C curve is below both vowel curves). The second option is when the highest point on the C curve is above both vowel curves. The third possibility is when the highest point on the C curve is above the V1 curve and below the V2 curve. The fourth case is when the highest point on the C curve is below the V1 curve and above the V2 curve.

The pattern where the highest point on the C curve is below both curves corresponding to V1 and V2, is called a "trough" in this study (see Figure 4-2 for an example).

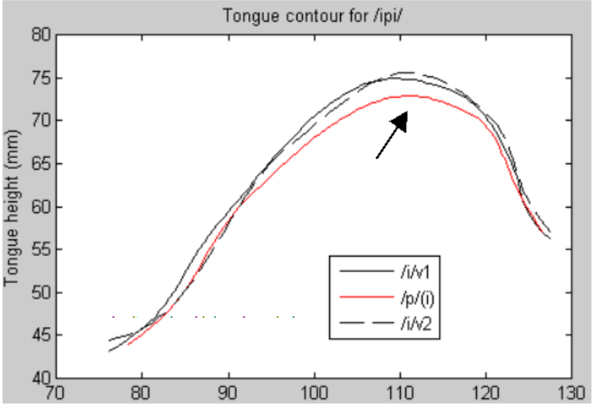


Figure 4-2. An example of the "trough" pattern, speaker S1. Solid black line – V1 curve; red line – C curve; dashed black line – V2 curve. The reference point on the C curve (at x = 111.20 mm and y = 72.94 mm) is indicated by an arrow.

In Figure 4-2, the reference point on the C curve is at x = 111.20 mm and y = 72.94 mm. The V1 curve is above this C curve reference point, showing that the tongue had moved down after the first vowel. The V2 curve is also above the reference point on the C curve, showing that the tongue had moved up after the consonant. Note also that the highest points on the vowel curves in this figure do not have the same x values as the highest point on the C curve (the highest point on the V1 curve is at x = 109.82 mm and y = 74.96 mm, and the highest point on the V2 curve is at x = 110.72 mm and y = 75.67 mm). This is typical of three-curve sequence patterns: the highest points on the vowel curves generally do not have the same x values as the highest point on the C curve.

The opposite pattern, where the highest point in the C curve is above both vowel curves, is called an “antitrough”. Figure 4-3 is an example. In this case, the reference point on the C curve is at x = 84.06 mm and y = 80.65 mm.

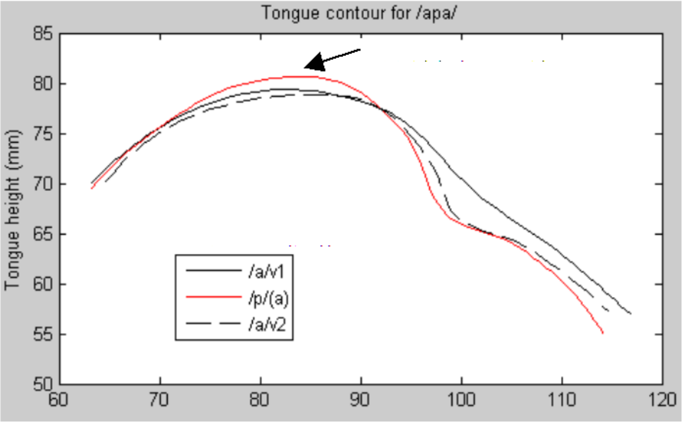


Figure 4-3. An example of the “antitrough” pattern, speaker S2. Solid black line – V1 curve; red line – C curve; dashed black line – V2 curve. The reference point on the C curve (at x = 84.06 mm and y = 80.65 mm) is indicated by an arrow.

The pattern where the highest point in the consonant contour is in the middle of the two vowel contours is illustrated in Figure 4-4. In this case, there occurs a continuous tongue movement upwards throughout the VCV. In this study, this pattern is called “continuous up”. The pattern where the tongue continuously moves downwards throughout the VCV is called “continuous down”.

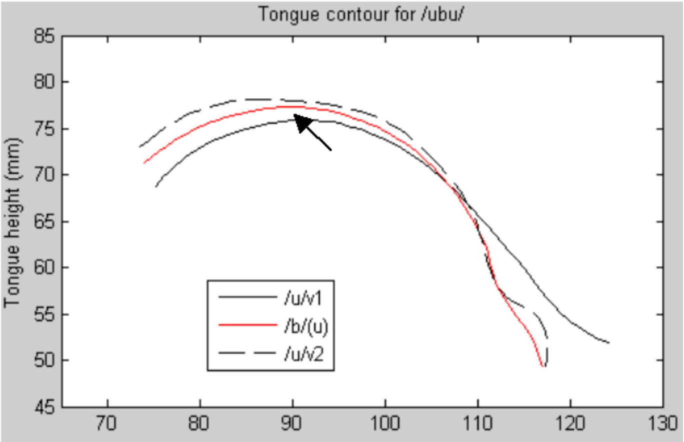


Figure 4-4. An example of the "continuous up" pattern, speaker S3. Solid black line – V1 curve; red line – C curve; dashed black line – V2 curve. The reference point on the C curve curve (at x = 90.11 mm and y = 77.26 mm) is indicated by an arrow.

The four tongue contour sequence patterns described above are summarised in Table 4-1. The numbers of troughs, antitroughs, continuous up and continuous down tongue contour sequence patterns were calculated and compared across subjects and across different vowel environments.

		Second movement (C to V2)	
		up	*down*
First movement (V1 to C)	*down*	Trough	Continuous down
	up	Continuous up	Antitrough

Table 4-1. Four types of tongue contour sequence patterns used in the analysis. Labels "up" and "down" indicate movement of the tongue up or down between V1 and C, and between C and V2.

4.2.7. Statistical analysis: binomial distribution

Statistical analysis of the numbers of troughs, antitroughs, continuous up and continuous down patterns was based on the binomial distribution. A significance threshold of 0.05 was used. As four separate tests were conducted for the four possible tongue contour sequence patterns, the Bonferroni adjustment was applied, so the cut-off threshold in each individual test was lowered to 0.0125. (The Bonferroni adjustment ensures that the overall risk for the four tests remains 0.05.) For a description of the binomial distribution, see Appendix IV.

4.3. Results

4.3.1. V-on-C coarticulation in bilabial consonants

In Figure 4-5, tongue contours of the same consonant in three different vowel environments are presented, for subject S1. It is obvious from the figure that the three consonant's shapes are different, depending on the vowel. This figure illustrates the pattern that was observed in all three subjects.

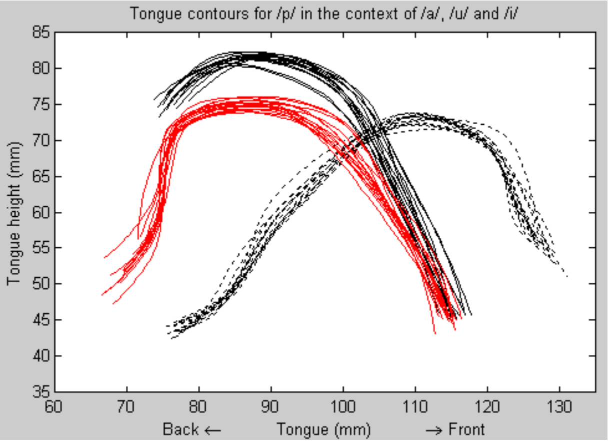
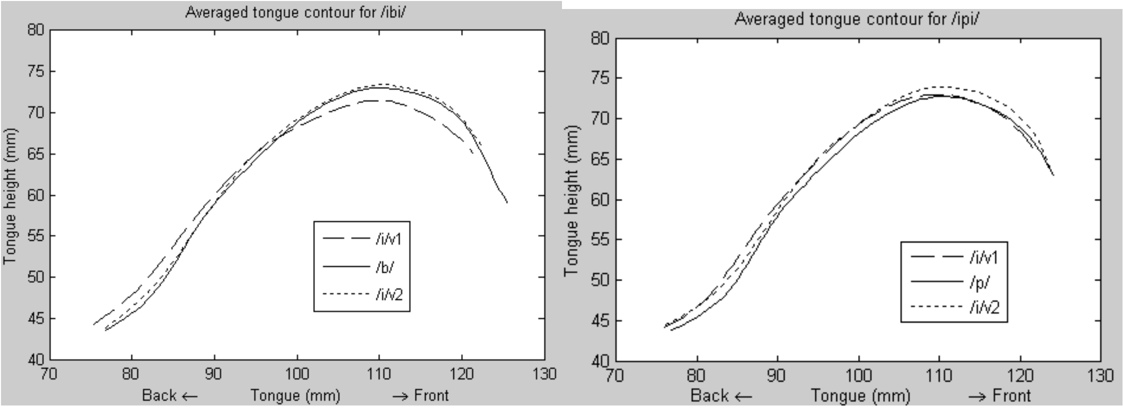


Figure 4-5. Tongue contours for the voiceless bilabial stop in the context of three different vowels, subject S1. Red lines – fifteen tokens of /p/ in the context of /a/; solid black lines – fifteen tokens of /p/ in the context of /u/; dashed black lines – fifteen tokens of /p^j/ in the context of /i/.

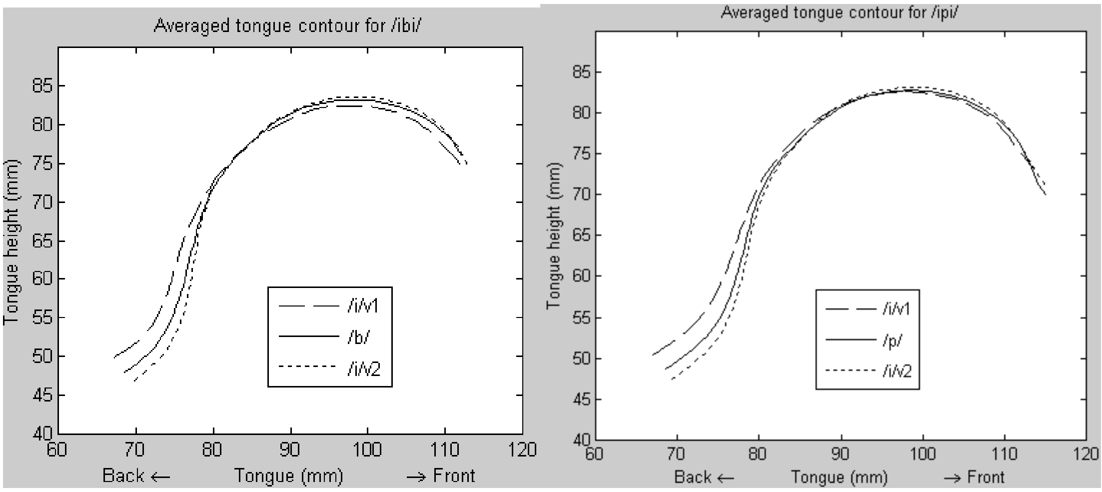
To obtain more data on coarticulation of bilabial consonants with surrounding vowels in Russian VCV sequences, let us look at plots where the curves for V1, C and V2 belonging to the same VCV sequence are displayed on the same graph. Figures 4-6 – 4-8 contain plots where the V1 curve, the C curve and the V2 curve are presented on one graph, each averaged over 15 tokens, by VCV type, for each subject.

Experiment 1: Russian bilabial stops

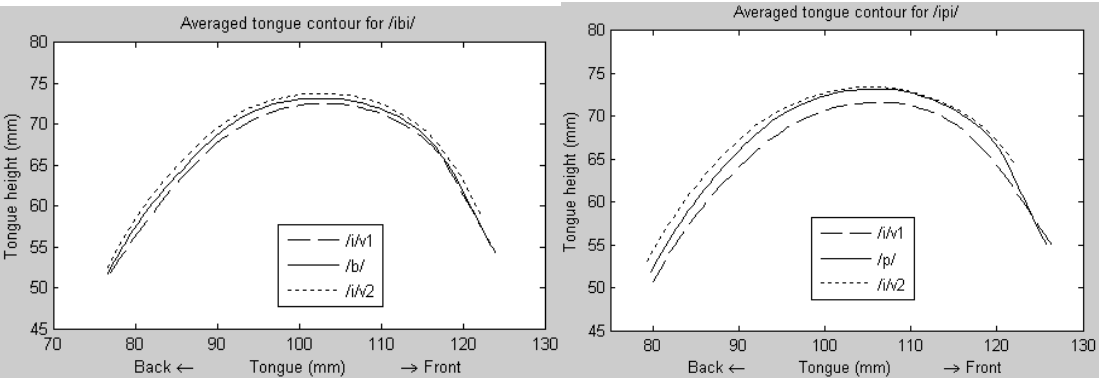
In Figure 4-6, average V1, C and V2 curves for /ibʲi/ and /ipʲi/ sequences are displayed.



a) S1, /ibʲi/ b) S1, /ipʲi/



c) S2, /ibʲi/ d) S2, /ipʲi/

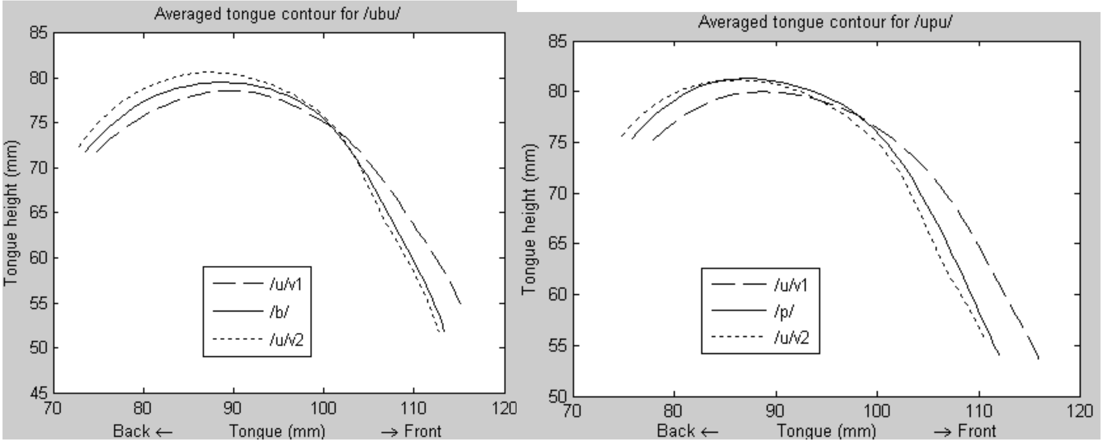


e) S3, /ibʲi/ f) S3, /ipʲi/

Figure 4-6. Average V1, C and V2 curves for /ibʲi/ and /ipʲi/ sequences, for the three subjects: a)-b) S1; c)-d) S2; e)-f) S3. Dashed line – V1 curve; solid line – C curve; dotted line – V2 curve.

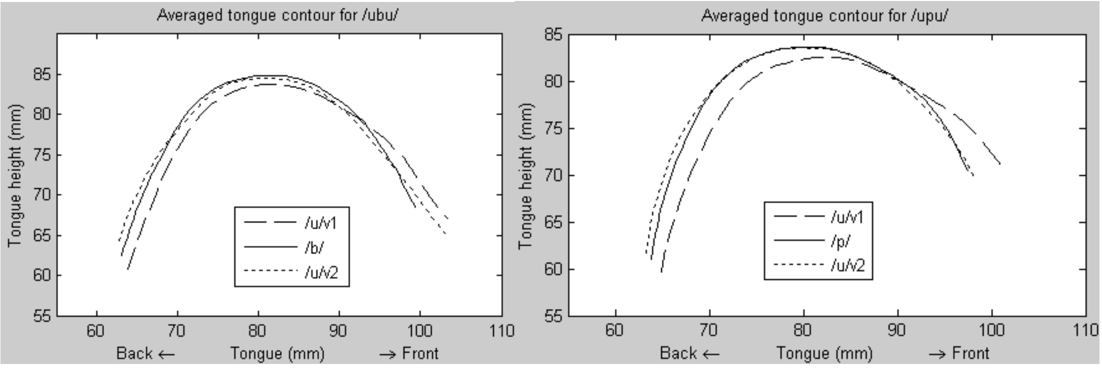
We notice that the C curve is very close to the vowel curves. This fact confirms the observation made about Figure 4-5, that the tongue shape for the consonant in a VCV sequence varies according to the surrounding vowels.

In Figure 4-7, average V1, C and V2 curves for /ubu/ and /upu/ sequences are displayed. In /uCu/ sequences, the three tongue curves displayed on each graph are also close together. A difference between /uCu/ and /iC^{j}i/ sequences is that in /uCu/, the tongue shapes for the three segments appear to be generally further apart than in /iC^{j}i/. Some reasons will be presented in Section 4.4.7.



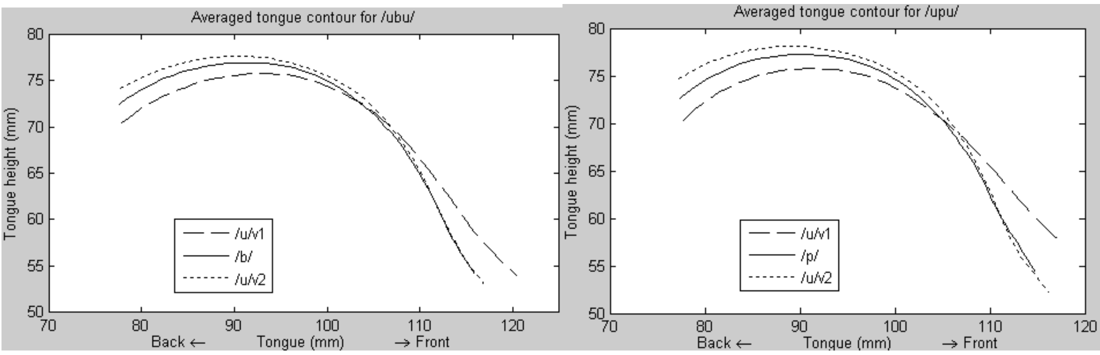
a) S1, /ubu/

b) S1, /upu/



c) S2, /ubu/

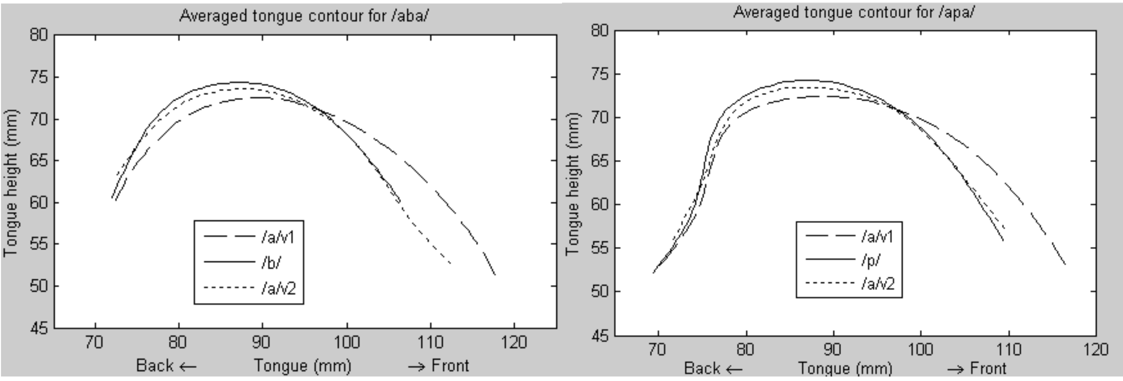
d) S2, /upu/



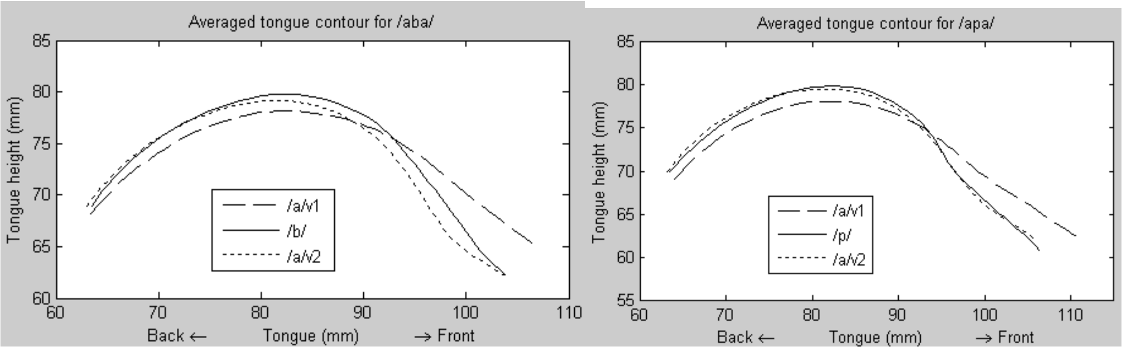
e) S3, /ubu/ f) S3, /upu/

Figure 4-7. Average V1, C and V2 curves for /ubu/ and /upu/ sequences, for the three subjects: a)-b) S1; c)-d) S2; e)-f) S3. Dashed line – V1 curve; solid line – C curve; dotted line – V2 curve.

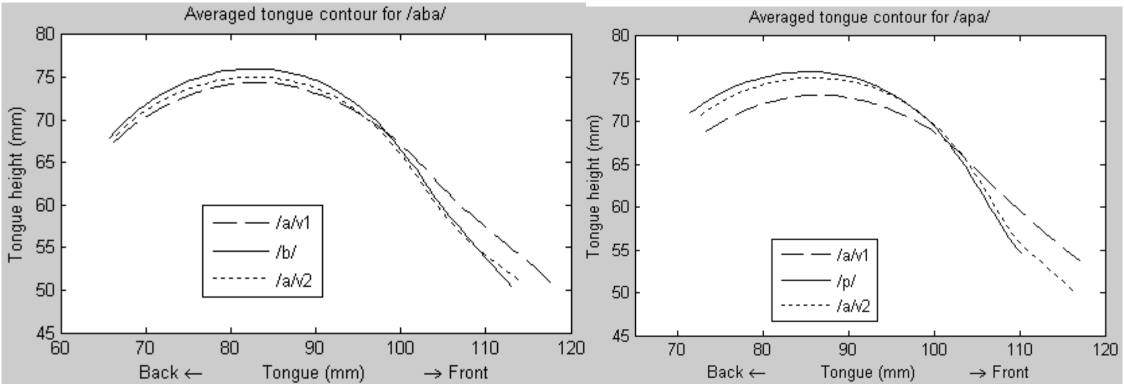
In Figure 4-8, average V1, C and V2 curves for /aba/ and /apa/ sequences are presented.



a) S1, /aba/ b) S1, /apa/



c) S2, /aba/ d) S2, /apa/



e) S3, /aba/ f) S3, /apa/

Figure 4-8. Average V1, C and V2 curves for /aba/ and /apa/ sequences, for the three subjects: a)-b) S1; c)-d) S2; e)-f) S3. Dashed line – V1 curve; solid line – C curve; dotted line – V2 curve.

In Figures 4-6 to 4-8, the three tongue curves displayed on each graph are rather close together, more so in /iC^{j}i/ than in /aCa/ or /uCu/. A common feature across all three subjects in /aCa/ sequences is that the tongue dorsum, or, to be more precise, the highest point of the tongue curve, is higher for the consonant than for both vowels. This observation is consistent with the results that will be presented in Sections 4.3.2 and 4.3.3, and discussed in Section 4.4.3.

The general observation about Figures 4-6 – 4-8 is that the contours of V1, C and V2 are rather close together in VCV sequences. This observation is consistent with the data presented in Figure 4-5, where the shape of the C curve was shown to be strongly differentiated according to the vowel environment. All these data show that the tongue surface shape during a bilabial consonant closure differs greatly according to the identity of the surrounding vowels.

4.3.2. Distribution of tongue contour sequence patterns

In this section, quantitative distribution of different tongue contour sequence patterns is analysed. In Section 4.2.6, we defined four theoretically possible patterns: “trough” (where the highest point in the C curve is below both vowel curves), “antitrough” (the highest point in the C curve is above both vowel curves), “continuous up” (the highest point in the C curve is above the V1 curve and below the V2 curve), and “continuous down" (the highest point in the C curve is below the V1 curve and above the V2 curve). In our data, only the first three of these patterns were present.

In Table 4-2, the numbers of occurrences of tongue contour sequence patterns are presented for all the subjects and all the VCV types pooled (270 repetitions in total). According to the binomial experiment conditions, in a case of 270 tokens with four alternatives, a pattern occurs significantly above chance level (at $p < 0.05$) if its number of occurrences is equal to or more than 85, and significantly below chance level if its number of occurrences is equal to or fewer than 51.

Trough	Antitrough	Continuous up	Continuous down
36	81	153	0

Table 4-2. Numbers of occurrences of tongue contour sequence patterns, for all the subjects together, and all the VCV types together.

Figure 4-9 illustrates the distribution of the three different patterns in all the speakers together. It shows that troughs occurred in 13% of the tokens, antitroughs in 30% of the tokens, and 57% had "continuous up" patterns.

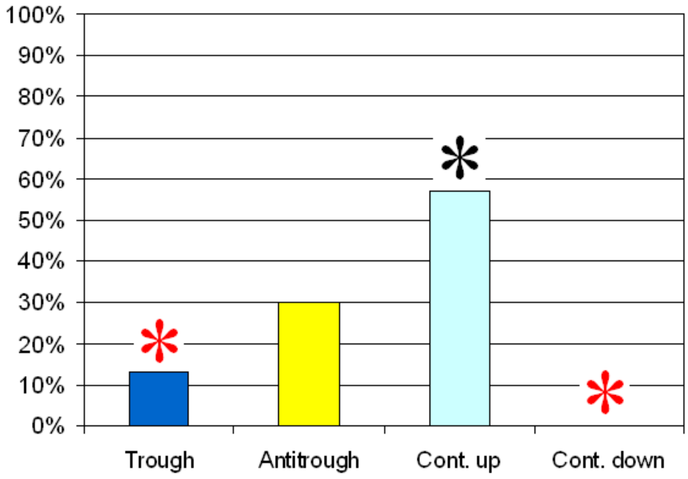


Figure 4-9. Distribution of tongue contour sequence patterns, for all the subjects together (percentage, out of 270 tokens). A black asterisk above a bar means that the rate of occurrence was significantly above chance. A red asterisk above a bar means that the rate of occurrence was significantly below chance.

The rate of occurrence of troughs and "continuous down" patterns was significantly below chance ($p < 0.001$ in both cases), and the rate of occurrence of "continuous up" was significantly above chance ($p < 0.001$), while the rate of occurrence of antitroughs was at chance level.

In Table 4-3, the numbers of occurrences of tongue contour sequence patterns are presented by vowel type, for all the subjects together. In each vowel group, there are 90 tokens. According to the binomial experiment conditions, in a case of 90 repetitions, a pattern occurs significantly above chance level if its number of occurrences is equal to or more than 33, and significantly below chance level if its number of occurrences is equal to or fewer than 13.

In Figure 4-10, the distribution of the patterns according to vowel type is presented.

	Trough	Antitrough	Continuous up	Continuous down
i	28	3	59	0
u	8	18	64	0
a	0	60	30	0

Table 4-3. Numbers of occurrences of tongue contour sequence patterns by vowel type, for all the subjects together.

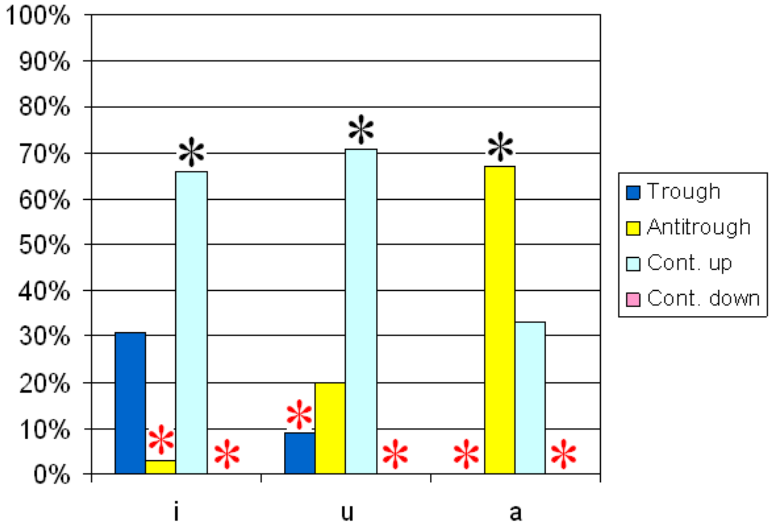


Figure 4-10. Distribution of tongue contour sequence patterns by vowel, for all the subjects together (percentage, out of 90 tokens). A black asterisk above a bar means that the rate of occurrence was significantly above chance. A red asterisk above a bar means that the rate of occurrence was significantly below chance.

We see that in the vowel /a/ context, the rate of antitrough occurrence is significantly greater than chance ($p < 0.001$). The rates of occurrence of troughs in /i/ and /u/ contexts are not significantly above chance, and in /u/, the rate of trough occurrence is even significantly below chance ($p < 0.001$). The rates of occurrence of antitroughs in /i/ and /u/ environments are also not significantly above chance: at chance level in /u/, and below chance in /i/ ($p < 0.001$). In contrast, continuous upward tongue movement occurs in these two vowel environments significantly above chance ($p < 0.001$ in both cases). In the /a/ context, the rate of occurrence of continuous patterns is at chance level.

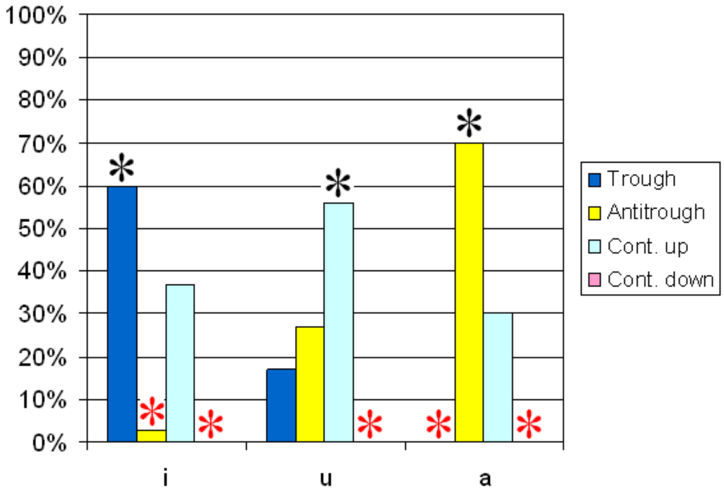
4.3.3. Distribution of tongue contour sequence patterns, individual results

Here, the data are analysed for the three subjects separately. In Table 4-4, the numbers of occurrences of tongue contour sequence patterns are presented according to subject and vowel type. In each group (represented by a row in the table), there are 30 tokens. According to the binomial experiment conditions, in a case of 30 repetitions, a pattern occurs significantly above chance level if its number of occurrences is equal to or more than 14, and significantly below chance level if its number of occurrences is equal to or fewer than 2.

		Trough	Antitrough	Continuous up	Continuous down
S1	i	18	1	11	0
	u	5	8	17	0
	a	0	21	9	0
S2	i	10	2	18	0
	u	3	10	17	0
	a	0	18	12	0
S3	i	0	0	30	0
	u	0	0	30	0
	a	0	21	9	0

Table 4-4. Numbers of occurrences of tongue contour sequence patterns, by subject and by vowel type.

In Figure 4-11, tongue contour sequence patterns are presented according to subject and vowel type. We see that in the /a/ context, the pattern is very similar across subjects: the rates of occurrence of antitroughs are significantly above chance in all the subjects ($p < 0.001$). In the two high vowels, there is considerable variability among subjects. In S1, the rate of occurrence of troughs in /iC^{j}i/ sequences is above chance ($p < 0.001$). In /uCu/ sequences, the rate of occurrence of troughs is at chance level in both S1 and S2. In the vowel /u/ context, the rates of occurrence of continuous tongue movements are significantly above chance in all the subjects ($p < 0.001$). Another fact adding to the cross-subject variability is the distribution of tongue contour sequence patterns in the context of high vowels in S3. In this subject, unlike the others, there are no troughs or antitroughs in these two vowels, only “continuous up” tongue movements throughout the VCVs.



a)

Experiment 1: Russian bilabial stops

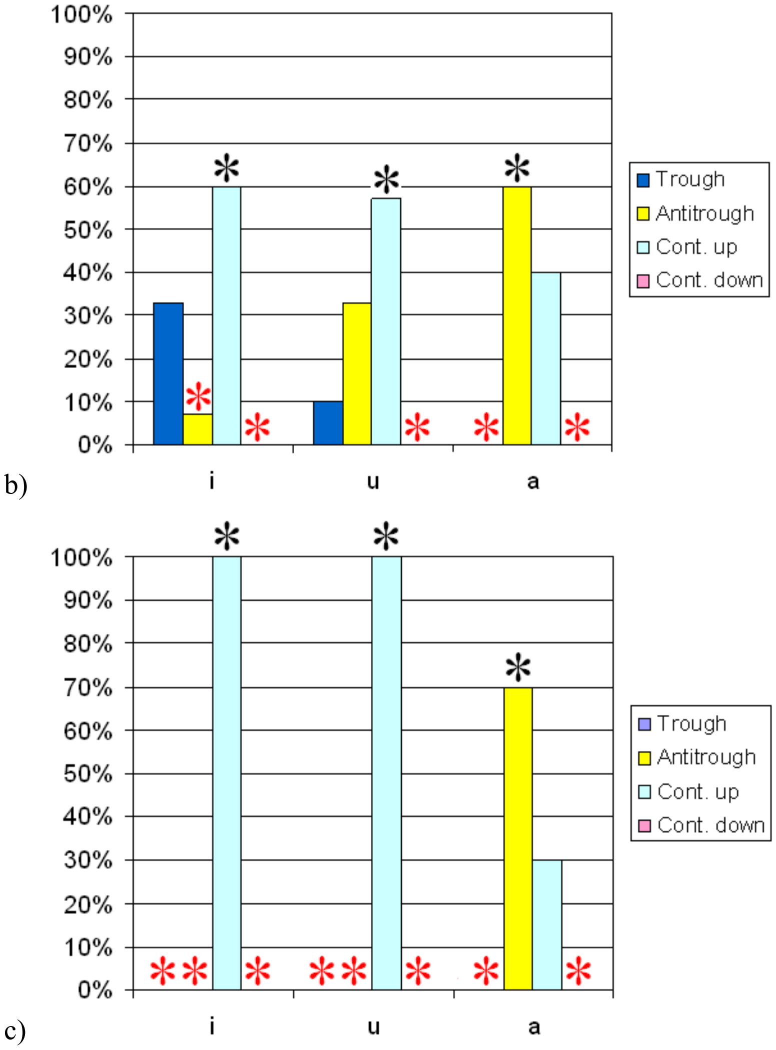


Figure 4-11. Distribution of tongue contour sequence patterns, by subject and by vowel type (percentage, out of 30 tokens): a) subject S1; b) subject S2; c) subject S3. A black asterisk above a bar means that the rate of occurrence was significantly above chance. A red asterisk above a bar means that the rate of occurrence was significantly below chance.

4.3.4. Distribution of tongue contour sequence patterns in voiceless versus voiced consonants

In Table 4-5, the numbers of occurrences of tongue contour sequence patterns are presented for all the subjects together, for voiceless and voiced consonants separately. In each group (represented by a row in the table), there are 45 tokens. According to the binomial experiment conditions, in a case of 45 repetitions, a pattern occurs significantly above chance level if its number of occurrences is equal to or more than 19, and significantly below chance level if its number of occurrences is equal to or fewer than 4 (at $p < 0.05$).

Figure 4-12 presents the distribution of the tongue contour sequence patterns in all the subjects together, for voiceless and voiced consonants separately.

When we look at the vowel /a/ environment, we notice that the rate of occurrence of antitroughs is significantly above chance in both /aba/ and /apa/ ($p < 0.001$), but it is greater in /aba/ than in /apa/. In both high vowel environments, “continuous up” patterns always occur at significantly greater than chance level ($p < 0.001$). As for the rate of trough occurrence, it is at chance level both in /ip^{j}i/ and in /ib^{j}i/. In /ubu/ and /upu/, the rate of occurrence of troughs is smaller than in in the /i/ context; it stays at chance level in /ubu/, and is below chance in /upu/ ($p < 0.001$). In /ubu/ sequences, antitroughs are fewer than chance ($p < 0.002$), and in /upu/ sequences, antitroughs are at chance level.

	Trough	Antitrough	Continuous up	Continuous down
ib^{j}i	11	2	32	0
ip^{j}i	17	1	27	0
ubu	6	3	36	0
upu	2	15	28	0
aba	0	37	8	0
apa	0	23	22	0

Table 4-5. Numbers of occurrences of tongue contour sequence patterns, for all the subjects together, for voiceless and voiced consonants separately.

Experiment 1: Russian bilabial stops

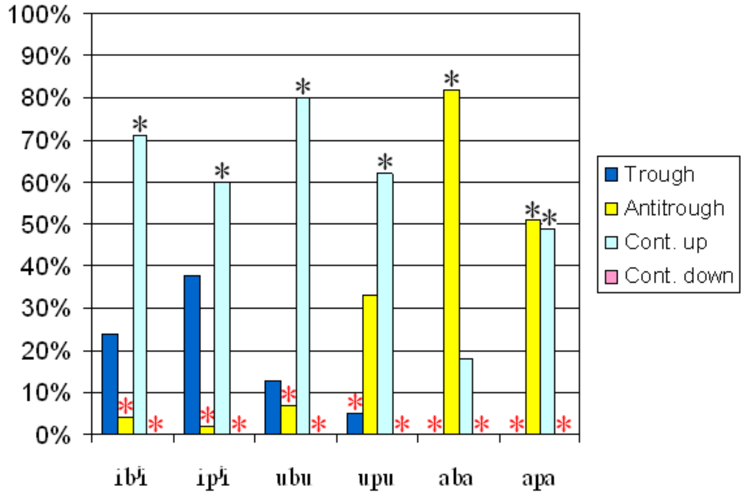


Figure 4-12. Distribution of tongue contour sequence patterns in voiced versus voiceless consonants, for all the subjects together (percentage, out of 45 tokens). A black asterisk above a bar means that the rate of occurrence was significantly above chance. A red asterisk above a bar means that the rate of occurrence was significantly below chance.

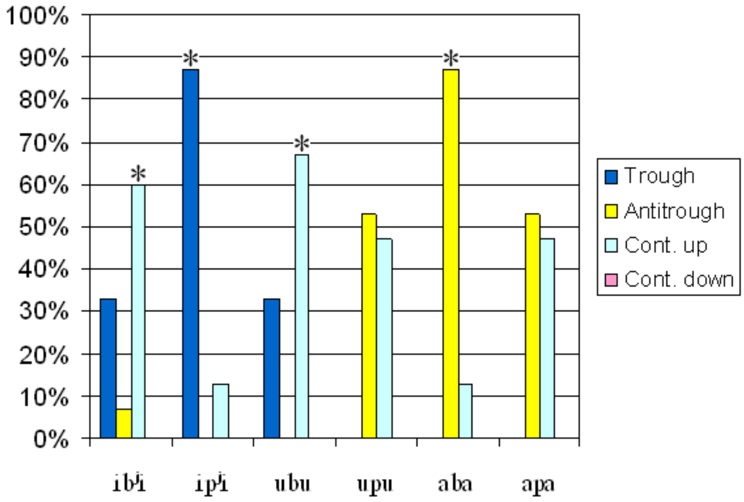
In Table 4-6, the numbers of occurrences of tongue contour sequence patterns are presented, by subject, by vowel type and by consonant type. In each group (represented by a row in the table), there are 15 tokens. According to the binomial experiment conditions, in a case of 15 repetitions, a pattern occurs significantly above chance level (at $p < 0.05$) if its number of occurrences is equal to or more than 9. In this case, no patterns can occur significantly below chance level, because zero occurrences has a mathematical probability of occurrence of 0.0134, which is above the threshold of 0.0125 resulting from the Bonferroni adjustment (see Section 4.2.7 and Appendix IV for more details).

In Figure 4-13, the patterns are presented according to subject, vowel type and consonant type. When looking at the /a/ context, we can see that in S1 and S2, there are more antitroughs in /aba/ than in /apa/. In /aba/, in both subjects the rate of occurrence of antitroughs is significantly above chance ($p < 0.001$), and in /apa/, in both subjects the rate of antitrough occurrence is at chance level. In S3, the difference in the pattern of distribution between the two /aCa/ sequences is in the same direction as in the other two subjects, but very small; in both sequences, the rate of occurrence of antitroughs is smaller than in S1 and S2, but it is above chance ($p < 0.001$) both in /aba/ and /apa/.

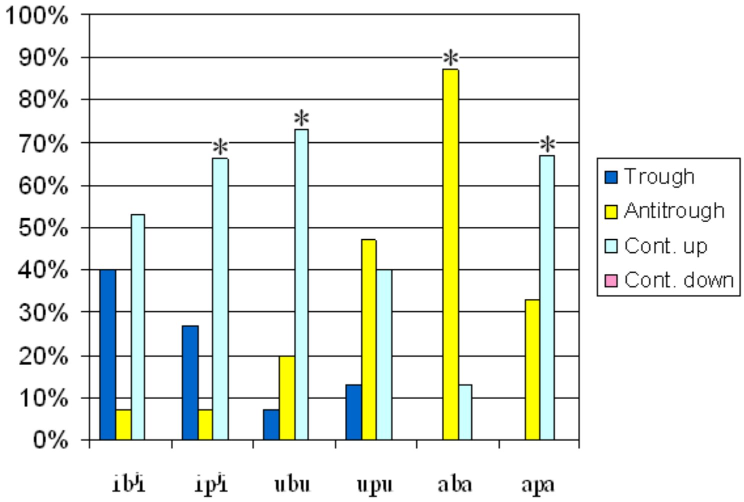
		Trough	Antitrough	Continuous up	Continuous down
S1	ib^{j}i	5	1	9	0
	ip^{j}i	13	0	2	0
	ubu	5	0	10	0
	upu	0	8	7	0
	aba	0	13	2	0
	apa	0	8	7	0
S2	ib^{j}i	6	1	8	0
	ip^{j}i	4	1	10	0
	ubu	1	3	11	0
	upu	2	7	6	0
	aba	0	13	2	0
	apa	0	5	10	0
S3	ib^{j}i	0	0	15	0
	ip^{j}i	0	0	15	0
	ubu	0	0	15	0
	upu	0	0	15	0
	aba	0	11	4	0
	apa	0	10	5	0

Table 4-6. Numbers of occurrences of tongue contour sequence patterns, by subject, by vowel type and by consonant type.

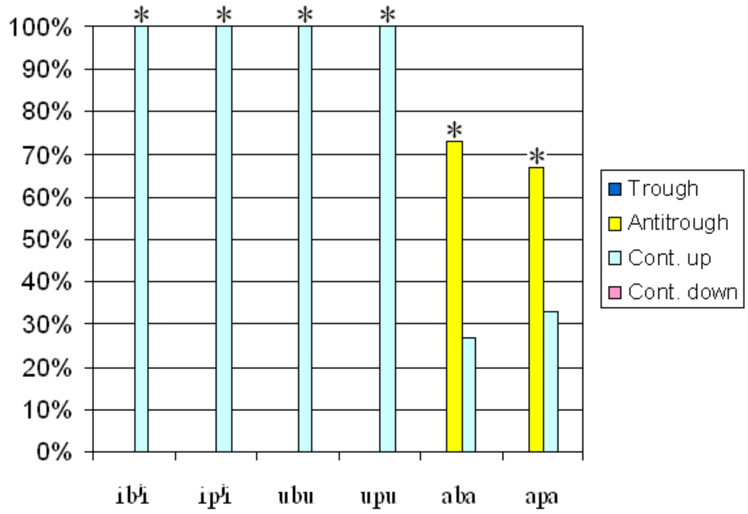
Experiment 1: Russian bilabial stops



a)



b)



c)

Figure 4-13. Distribution of tongue contour sequence patterns by subject, by vowel type and by consonant type (percentage, out of 15 tokens): a) subject S1; b) subject S2; c) subject S3. A black asterisk above a bar means that the rate of occurrence was significantly above chance.

For the vowel /i/, Figure 4-13 shows that the only VCV sequence that has a significantly higher than chance rate of trough occurrence, is /ip^{j}i/ in S1. In /ib^{j}i/, the distribution is similar in S1 and S2: in S1, “continuous up” patterns occur at greater than chance level ($p < 0.01$); in S2, “continuous up” patterns are at chance level. In /ip^{j}i/ in S2, the distribution is similar to /ib^{j}i/ in this subject, only the rate of occurrence of “continuous up” patterns is significantly above chance in /ip^{j}i/ sequences, while it is at chance level in /ib^{j}i/ sequences. In /ubu/, in both S1 and S2, only continuous patterns are above chance ($p < 0.001$), and in /upu/, in both subjects all the patterns are at chance level. The distribution of the patterns in high vowels in S3, as we saw in Figure 4-11, is radically different from that of the other two subjects, in that S3 only has continuous upward movement from V1 to V2 in these vowel environments.

4.3.5. Stress influence on V1 and V2 curves

In Figures 4-6 – 4-8, we can notice that the shapes of the V1 curve and the V2 curve are different. The second vowel was stressed in our stimuli. Stress is the principal factor contributing to the difference in the shapes of V1 and V2 curves.

In Figures 4-6 – 4-7, where VCVs with high vowels are displayed, we notice that there is a continuous tongue movement from V1 to V2, throughout the consonant, and that the dorsum of the tongue is higher in the V2 curve than in the V1 curve. This is evidenced by the fact that the maximum y value of the V2 curve is greater than the maximum y value of the V1 curve.

In Figure 4-8 (/aCa/ sequences), the V1 curve is further forward than the V2 curve. This is evidenced by the fact that the back part of the tongue in the V1 curve is further to the right along the x axis than in the V2 curve. (The back part of the tongue is called "tongue root" here, even though not the whole root of the tongue was captured in the spline drawing, but only the part of it that was visible above the hyoid bone shadow.) The front third (the blade) of the tongue is also in a more fronted position in the V1 curve than in the V2 curve.

Figure 4-14 is an example of a typical pattern of the V1 curve in relation to the V2 curve, in an /ubu/ token.

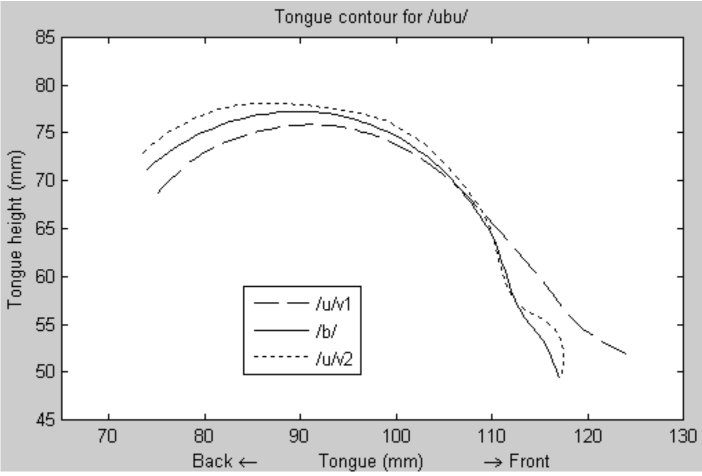


Figure 4-14. A three-curve pattern of a single token of /ubu/, produced by one subject (S3). Dashed line – V1 curve; solid line – C curve; dotted line – V2 curve.

4.3.6. Syllable boundary influence on VCV coarticulation

When we look at Figures 4-6 – 4-8 (Section 4.3.1), we notice that the C curve appears to be more similar to the V2 curve than to the V1 curve. In /iC^ji/ sequences, this is particularly obvious in /p^{j}i/ in S2 and S3 (Figure 4-5d and Figure 4-5f, respectively) and in /ib^{j}i/ in S1 and S2 (Figure 4-5a and Figure 4-5c, respectively). In /uCu/ sequences (Figure 4-6), there is a clearly visible pattern of the C curve being closer to the V2 curve than to the V1 curve. In /aCa/ sequences, the V1 curve is also noticeably different from the other two curves, across subjects.

Figure 4-15 shows an example of a typical pattern with the consonant being closer to the second vowel than to the first vowel.

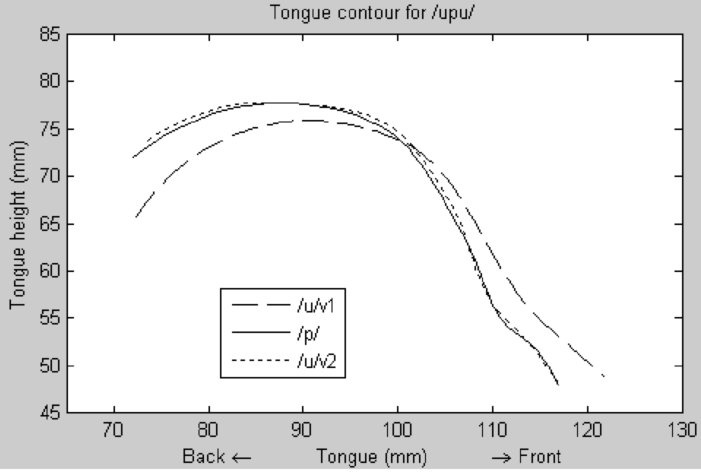


Figure 4-15. A three-curve pattern of a single token of /upu/, produced by one subject (S3). Dashed line – V1 curve; solid line – C curve; dotted line – V2 curve. Note that the C curve is very close to the V2 curve.

One more fact worth mentioning is that in /uCu/ and /aCa/ sequences, the front third (the blade) of the tongue is in a more fronted position in the V1 curve than in the V2 curve, as evidenced by the fact that the tongue blade in the V1 curve is further to the right along the x axis than in the V2 curve (see Figures 4-7 and 4-8). In /iC^{j}i/ sequences, we do not observe this pattern.

4.4. Discussion

4.4.1. V-on-C coarticulation in bilabial consonants

It was demonstrated in this experiment that the tongue shape for the bilabial consonant varies greatly, according to the adjacent vowels. Also, it was shown that within a VCV sequence, the tongue contours for V1, C and V2 are very close together. All these findings support Hypothesis 1. These are expected results, given the existing literature on lingual coarticulation in labial consonants (e.g., Recasens 1999), where it has been claimed that vowel-dependent lingual coarticulation is large in labial stops, across languages.

In terms of the DAC model, our findings demonstrate that the degree of resistance to vocalic coarticulation in bilabial stop consonants in Russian VCV sequences is lower than maximal.

The results of the qualitative analysis of V-on-C coarticulation in bilabial consonants demonstrated the need for a quantitative procedure for comparing sets of curves for significant differences, in order to be able to claim that the tongue shapes for a particular sound in different contexts differ significantly from each other.

4.4.2. Tongue movements in /iC^{j}i/ and /uCu/ sequences

The hypothesis that there would be a continuous tongue movement between two high vowels (Hypothesis 2) was generally supported. In both /i/ and /u/ contexts, the rate of occurrence of "continuous up" patterns was significantly above chance. In the /u/ environment, the rate of occurrence of troughs – tongue contour sequence patterns with a lowering between the two vowels – was at chance level or below chance in individual results for the three subjects, and below chance in the results for the three subjects pooled. In the /i/ context, in all the subjects pooled, the rate of occurrence of troughs was at chance level. In the individual results for /iC^{j}i/ sequences, one subject did not exhibit troughs at all, another subject had a chance level rate of trough occurrence, and only in one subject the rate of occurrence of troughs was significantly above chance. The rate of occurrence of antitroughs – tongue contour sequence patterns with a raising between the two vowels – in the /i/ context was below chance in all the subjects, individually and pooled. In the /u/ environment, antitroughs were at chance level in all the subjects together; in one subject, the rate of antitrough occurrence was below chance, and in two other subjects, antitroughs were at chance level.

So there is a lack of support from our Russian data for the discontinuity in coarticulation found in some languages (see Section 2.2.12). We may then claim that our data demonstrate language-specificity of this pattern of discontinuity in coarticulation in VCV sequences with bilabial consonants. Possible reasons for this result are given below.

Considering the characteristics of the Russian phonological system, described in Section 4.1.1, it is reasonable to suggest that phonological palatalisation and phonetic velarisation and pharyngealisation of Russian consonants constitute a contributing factor to the continuity in coarticulation found in Russian VCVs with high vowels. A constraint on tongue position for producing non-palatalised and palatalised consonants makes these Russian labial consonants more similar in tongue shape to the neighbouring high vowels than, for example, English labial consonants. This interpretation is reminiscent of Öhman (1966), where articulatory constraints on Russian consonant production were suggested to be the reason for the small freedom of coarticulation found in these consonants (see Section 2.3.2.2).

In terms of the DAC model, the absence of troughs in Russian can be interpreted as a manifestation of coarticulation resistance by the consonants which have some tongue position specification, largely compatible with the lingual gestures for the neighbouring vowels: namely, the raised and fronted tongue dorsum in palatalised consonants, and the raised and backed tongue dorsum in non-palatalised consonants.

A possible explanation of troughs in VCV sequences with labial consonants in some other languages within the DAC model will be offered in Section 4.4.5, where our data will be compared with the British English data presented in Vazquez Alvarez et al. (2004).

4.4.3. Tongue movements in /aCa/ sequences

The hypothesis which stated that the tongue would raise between two low vowels (Hypothesis 3) was supported, across subjects. These findings contradict the data presented in Svirsky et al. (1997). That study found consistent tongue *lowering* between the two vowels in American English /aCa/ sequences with bilabial consonants /b/ and /p/. Tongue dorsum displacement was measured in that study using EMA. Svirsky et al. (1997) suggested that the discontinuity in lingual coarticulation occurred because of the intraoral pressure build up, pushing the tongue downwards between the vowels.

Our findings are partly consistent with the British English data presented in Vazquez Alvarez et al. (2004). In that study, both tongue raising and tongue lowering were found in British English /aCa/ sequences with bilabial stop consonants. The researchers suggested that the observed intervocalic tongue raising could be a manifestation of the tongue returning to the neutral, schwa-like position.

The following explanation of our results appears logical, consistent with other results in this work, and also accounting for the differences between our results and the results reported in other published studies.

As well as in the case of /iC^{j}i/ and /uCu/ sequences, the observed tongue movement patterns in /aCa/ sequences could be explained by the characteristics of the Russian phonological system. In /aCa/ sequences, the consonants investigated in this work were phonologically non-palatalised (as opposed to palatalised). In Section 4.1.1, Russian non-palatalised consonants are described as phonetically velarised and pharyngealised. This velarisation and pharyngealisation could account for the tongue raising between the two low vowels. A constraint on the tongue position during production of phonetically velarised and pharyngealised non-palatalised consonants would result in significant tongue raising in /aCa/ sequences in Russian, not observed in, e.g., English /aCa/ sequences. Qualitative observations of whole contours confirm these suggestions. In Figure 4-8, the tongue dorsum (i.e., the middle third of the tongue contour) in the C curve is generally higher than in the two vowel curves; the tongue root is further backward than the V1 curve, and further backward or very close to the V2 curve.

In terms of the DAC model, antitroughs in Russian /aCa/ sequences with labial consonants could be considered a demonstration of CR by non-palatalised consonants. These consonants have a tongue position specification, which is not compatible with the lingual gesture for the neighbouring low vowels: namely, the raised tongue dorsum and backed tongue root. Tongue displacement upwards and backwards would be interpreted as a manifestation of resistance to coarticulation by bilabial stop consonants.

Some experimental data from the literature suggest that jaw position may affect tongue movements in /aCa/ sequences with bilabial consonants. For example, Vazquez Alvarez et al. (2004) observed some intervocalic tongue dorsum raising in these sequences, and claimed that jaw raising could have possibly contributed to these patterns. Fuchs et al. (2004), in an EMA experiment, observed intervocalic tongue raising in /apa/ and /aba/ sequences (tongue blade displacement was reported in that study). In a different condition, involving a bite block, the raising was smaller, and even some tongue lowering was observed in /apa/ sequences. So in the present work, jaw influence cannot be discarded in the case of bilabial stop consonants.

4.4.4. Voiced versus voiceless consonants

The hypothesis that there would be no consistent differences in lingual coarticulation between voiced and voiceless consonants (Hypothesis 4) was supported by our data. In Section 4.3.1, we saw that the plots with average V1, C and V2 curves look very similar in VCVs with voiced and voiceless consonants. The analysis of the tongue contour sequence pattern distribution also produced similar patterns for VbV and VpV sequences. All these results suggest that the tongue movements in our data did not systematically contribute to creating the distinction between voiced and voiceless consonants.

This finding is in contradiction with some results reported in the literature. For example, Svirsky et al. (1997) reported significantly greater tongue lowering in /aba/ sequences than in /apa/ sequences in American English. Svirsky et al. (1997) explained that pattern by active tongue displacement during /b/, in order to accommodate airflow into the oral cavity, for maintaining vocal fold vibration. Fuchs et al. (2004) studied

German VCVs with high vowels (/i/ and /u/) using EMG, and observed significantly greater tongue lowering in /VpV/ sequences than in /VbV/ and /VmV/ sequences. Fuchs et al. (2004) suggested that duration of the intervocalic consonant could contribute to producing troughs: /b/ and /m/ in their data were consistently shorter than /p/.

In this work, lingual coarticulatory patterns in VCV sequences with voiced and voiceless consonants were as predicted by Hypothesis 2 (for /iC^{j}i/ and /uCu/ sequences) and Hypothesis 3 (for /aCa/ sequences). Similar lingual coarticulatory patterns in VCVs with voiced and voiceless consonants in our data could be explained by the fact that in Russian, unlike in English or German, bilabial consonants are constrained in lingual position for their production (as described in Section 4.1.1 and discussed in Sections 4.4.2 and 4.4.3). This phonologically based constraint on lingual position in Russian bilabial consonants probably overrules any aerodynamic or physiological requirements on the tongue during production of voiced and voiceless stops.

An interesting cross-subject difference was observed in these data, in respect of the tongue behaviour in VCVs with voiced versus voiceless consonants. Two subjects, S1 and S2, demonstrated the following lingual coarticulatory patterns in /aba/ and /apa/ sequences: the rate of occurrence of antitroughs was significantly above chance level in /aba/, but it was at chance level in /apa/. A plausible explanation for these patterns could be as follows.

Kozhevnikov and Chistovich (1965) studied the displacement of the jaw and the lower lip in /apa/, /aba/ and /ama/ sequences in Russian. The physiological mechanisms described by these researchers account very well for our results on voiced versus voiceless consonants. In our data, the greater number of antitroughs in /aba/ than in /apa/ could have been due to the differences in the timing of the jaw and the lower lip displacement in VCV sequences. According to Kozhevnikov and Chistovich (1965), the jaw raises faster in /aba/ than in /apa/ sequences, while the lower lip moves slower in /aba/ than in /apa/. The tongue, depending on the jaw, then has to make a greater movement in /aba/ than in /apa/. In our data, this would account for a greater number of antitroughs in /aba/ than in /apa/ sequences.

Some support for this explanation of the difference between aCa sequences with voiced and voiceless consonants comes from the temporal model of speech production (Bell-Berti and Harris 1981; see Section 2.3.2.1 for description). One rule for the timing of articulatory activity underlying segment representation, postulated in the model, claims that for the production of each speech sound, “the articulatory period may begin at different times for different articulators” (Berti & Harris 1981, p. 16). This rule describes very well the trade-off among the three articulators involved in the production of bilabial stops surrounded by two identical low vowels.

In subject S3, this pattern was not observed: S3 had a similar number of antitroughs in /aba/ and /apa/ sequences. This cross-subject difference could suggest that timing relationships of different articulators for production of voiced and voiceless consonants in Russian may vary across speakers, as long as the consonants are correctly perceived. However, examination of whole curves suggests that this pattern in S3’s results could have been due to transducer orientation under the chin. In this subject, the transducer was slightly rotated clockwise, and this may have affected the results of the calculations based on the vertical measure bar. Had the transducer been at a right angle to the jaw, the calculations for /aba/ and /apa/ sequences in S3 would have produced a pattern more similar to that found in the other two subjects (see Section 4.4.9 for details of how transducer orientation may affect the results of such calculations).

4.4.5. Cross-linguistic comparison of tongue movements in VCV sequences with non-lingual consonants

Here, the results of Experiment 1 are discussed in relation to the results of the ultrasound study presented in Vazquez Alvarez et al. (2004). In that study, there were ten British English subjects, and each subject produced five repetitions of six different VCV sequences: /ibi/, /ipi/, /ubu/, /upu/, /aba/, /apa/. The subjects were asked to produce the target VCVs with equal stress on both syllables. Recording and analysis were conducted using the QMUC ultrasound system (see Section 3.3). The method used in Experiment 1 in the present work, for identifying and calculating the numbers of different tongue contour sequence patterns, was also used in the study by Vazquez Alvarez and her colleagues (see the description of the method in Section 4.2.6). In this section, two sets of results are compared: Russian VCVs with bilabial stops (Experiment 1) and British English VCVs with bilabial stops (Vazquez Alvarez et al. 2004). Vazquez Alvarez et al. (2004) did not use the binomial distribution in their statistical analysis. So in the present study, statistical calculations were performed on the data presented in Vazquez Alvarez et al. (2004), in order to obtain a comparable representation of the results from that study with the present work.

The distribution of the four tongue contour sequence patterns in Russian and English data is presented in Table 4-7. In the Russian set of data, there were 90 tokens in each of the three vowel groups, and in the English set of data, there were 100 tokens in each vowel group. Therefore, raw numbers of occurrences are not given in the table, but percentage values, in order to facilitate comparison of the two sets of data.

	/i/		/u/		/a/	
	Russian	*English*	*Russian*	*English*	*Russian*	*English*
Trough	31	86**	9**	76**	0**	33
Antitrough	3**	3**	20	1**	67**	34
Continuous up	66**	5**	71**	0**	33	19
Continuous down	0**	6**	0**	23	0**	14*

Table 4-7. Rate of occurrence (percentage) of tongue contour sequence patterns, in Russian and British English VCV sequences with bilabial consonants. The data are presented according to vowel type, for all the subjects together. Two black asterisks next to a number mean that the rate of occurrence was significantly above chance, at $p < 0.001$. Two red asterisks next to a number mean that the rate of occurrence was significantly below chance, at $p < 0.001$. One red asterisk next to a number means that the rate of occurrence was significantly below chance, at $p < 0.01$. For details about statistical calculations, see Appendix IV.

We can see from the table that in English, the rate of occurrence of troughs is significantly above chance in /i/ and /u/ contexts ($p < 0.001$), while in Russian, it is at chance level or below.

In the low vowel environment, the distribution of tongue contour sequence patterns is again different in Russian and English data. In Russian, the rate of occurrence of antitroughs is significantly above chance level ($p < 0.001$), and there are no troughs. In the British English bilabial stops, the numbers of troughs and antitroughs are comparable, and both are at chance level.

When it is claimed that bilabial stops have a certain degree of resistance to lingual coarticulation, it is meant that these phonemes retain their own identity with respect to lingual position, to a certain degree. The DAC model would interpret discontinuity in tongue movement between the two vowels of a VCV sequence by claiming that the intervocalic consonant has a degree of CR which is higher than zero. Continuity would not necessarily mean lack of resistance, but could be due to comparable articulatory trajectories of the intervocalic consonant and the surrounding vowels.

Russian bilabial stops (both non-palatalised and palatalised), as was described in Section 4.1.1, have a raised tongue posture, as compared with the neutral tongue position. In the Russian ultrasound data in this experiment, there is no tongue lowering between the two high vowels, and there is tongue raising between the two low vowels. The absence of troughs in Russian in the context of high vowels can be interpreted as a manifestation of coarticulation resistance by the consonants having a lingual position specification comparable to that of the surrounding vowels (see Section 4.4.2). The significant number of antitroughs in /aCa/ sequences can be interpreted as a manifestation of CR by the consonants having a lingual position specification different from that of the surrounding vowels (see Section 4.4.3).

In English, unlike Russian, a significant number of troughs is observed in /iCi/ and /uCu/ sequences. In /aCa/ sequences, again unlike Russian, no patten occurs significantly above chance. We could assume that the lingual position identity of bilabial stops is described as a slightly neutralised tongue posture, as compared with the tongue position required for the high vowel production. Then we could claim that the tongue lowers between the two high vowels towards the posture of the bilabial stop identity. Thus, in terms of the DAC model, troughs in VCV sequences with labial consonants in English can be considered a demonstration of a higher than zero CR by the consonants. This interpretation would be reminiscent of the claims made in, e.g., Perkell (1986) and Fuchs et al. (2004), that bilabial stops are not completely unspecified for lingual position. As for the low vowel context, the number of tongue raising patterns is not significantly above chance probably because the target lingual position for the English bilabial stops is not as high as for Russian bilabial stops.

4.4.6. Stress influence on V1 and V2 curves

In Section 4.3.5, we observed that the two vowel curves, V1 curve and V2 curve, consistently differed in shape, in all three vowels. These differences in tongue shape could be explained by stress influence.

In the context of the two high vowels, qualitative observations of whole tongue contours revealed patterns of continuous tongue movement upwards from V1 to V2, throughout the consonant. This is evidence that in /iC^{j}i/ and /uCu/ sequences, for producing V2, the tongue consistently had a more peripheral position, in relation to the more neutralised V1. Large numbers of occurrences of the “continuous up” tongue contour sequence pattern (see Section 4.3.3, Figure 4-11) indicate the same tendency: the tongue moving from a more neutral position in V1 towards the more peripheral position in V2. These data support Hypothesis 5.

This visual impression is confirmed by the auditory-based transcription of the data. Across subjects, the second /i/ of /iC^{j}i/ is perceived by myself as a closed front vowel [i], and the first /i/ is perceived as an [ɪ]-like sound. The second /u/ of /uCu/ sequences is transcribed as a closed back vowel [u], while the first /u/ can be transcribed as a more open and fronted, [ʊ]-like vowel.

The pattern described here corresponds to the data described in the literature (see Section 4.1.1): stressed high vowels in Russian are more closed than unstressed high vowels, because of the quantitative and qualitative reduction of unstressed vowels.

In the vowel /a/ context, as well as in the two high vowel contexts, the shapes of V1 and V2 curves were different: the V1 curve was further forward than the V2 curve. The auditory-based transcription of /aCa/ sequences confirms these visual impressions. Across subjects, the first /a/ is perceived as an [ɛ]- or [ɜ]-like vowel, while the second /a/ is transcribed as an [a]- or [ɐ]-like vowel. These data support Hypothesis 6.

The results on /aCa/ sequences also correspond to the results from the literature on qualitative reduction of the unstressed vowels in Russian (see Section 4.1.1).

Interpreting these results within the DAC model, we can say that the DAC values of the vowels depend on whether they are stressed or unstressed. Within the DAC model, when a phoneme possesses a certain extent of resistance to coarticulation, it means that a phoneme retains (to that extent) its own identity. Qualitative reduction of unstressed vowels is a demonstration of these vowels losing their identity to a certain extent, under the influence of stress. So we can claim that the DAC value of the unstressed vowel (V1 in this experiment) is lower than the DAC value of the stressed vowel (V2 in this experiment), to the extent to which the V1 curve is different from the V2 curve.

4.4.7. Syllable boundary influence on VCV coarticulation

Another result of this experiment is that in V#CV sequences, the C curve was generally closer to the V2 curve than to the V1 curve, as demonstrated by visual observation. This can be interpreted as a syllable boundary influence on the VCV coarticulatory pattern, in line with previous research that has demonstrated such influence (e.g., Kozhevnikov and Chistovich 1965; Bondarko 1969; Gay 1977; Lindblom et al. 2002; also, see Section 4.1.1). These results support Hypothesis 7.

The DAC model would treat these data as evidence that the DAC value of the intervocalic consonant in V#CV sequences is influenced by the syllable affiliation of the consonant. In relation to V2, the DAC value of the consonant is smaller than in relation to V1. Thus, in order to adequately describe the influence of the syllable boundary on coarticulation in VCV sequences, the intervocalic consonant needs to be represented by two separate DAC values, one in relation to each vowel. In Section 7.3.3, this principle is used for quantifying CR of intervocalic consonants, based on the whole curve ultrasound data.

Another indication of syllable boundary influence on VCV coarticulation is the following. In /uCu/ and /aCa/ sequences, the front third (the blade) of the tongue is in a more fronted position in the V1 curve than in the V2 curve (see Figures 4-6 – 4-8). This position of the blade can partly be attributed to the influence on the first vowel from the preceding /t/ (because the preceding word ended in /t/), making the V1 more fronted than the V2. The lack of any comparable influence of the following /t/ on the V2 can be explained by the fact that in Russian, coarticulation in CV sequences is stronger than in VC sequences, even across word boundaries (see Section 4.1.1. for the description and references).

In /iC^{j}i/ sequences, the pattern of tongue blade coarticulation described in the previous paragraph is not observed. This difference between /i/ and the other two vowel contexts could be explained as follows. The left and right context for /iC^{j}i/ sequences in our data was chosen to be the palatalised consonant /t^j/. The articulation of this consonant involves tongue dorsum raising to the hard palate (together with the blade raising), in order to create the secondary articulation of palatalisation (see Section 4.1.1). Consequently, the tongue shape during the production of this consonant is very close to the tongue shape required for the vowel /i/ production. In /uCu/ and /aCa/ sequences, the left and right context for the VCV was the non-palatalised consonant /t/. In order to produce the /t/ occlusion, the tongue blade had to travel from the relatively lowered position required for the vowel production towards the upper teeth. Hence, in /uCu/ and /aCa/ sequences we observe a difference in tongue blade position between the V curve and the other two curves, and in /iC^{j}i/ sequences we do not observe such a difference.

4.4.8. Inter-subject differences

Average tongue curves presented in Section 4.3.1 demonstrated inter-speaker variability in tongue shapes. This variability can be partly attributed to the vocal tract morphology in individual speakers, partly to the exact position and and angle of the transducer under the chin, and partly to varying degree of visibility of the tongue contour in different speech sounds.

Some inter-subject differences in whole tongue curves cannot be attributed to individual tongue shapes, or to the orientation of the transducer under the chin. For example, in /iC^{j}i/ sequences, tongue movement from V1 to V2 took slightly different forms across subjects. This difference in tongue behaviour could be attributed to slightly differing strategies of vowel production in the three subjects. In Figure 4-6, we see that in S2, the tongue root is more advanced in the stressed V2 than in the unstressed V1. In S3, the whole V2 curve is above the V1 curve. In S1, the front third of the V2 curve is above the V1 curve. This inter-subject difference suggests that all these strategies for creating a difference in quality between the stressed vowel and the unstressed one are acceptable in Russian, i.e., stress is correctly perceived.

There are other inter-subject differences that cannot be explained by vocal tract morphology or transducer placement under the chin. In Section 4.3.3, some cross-subject differences were noticed in the distribution of tongue contour sequence patterns. For example, subject S3 was very different from the other two subjects. Namely, there were no troughs in that subject's results. Also, in the two high vowel contexts, S3 only produced "continuous up" patterns, unlike the other subjects. This tendency towards continuous tongue movement from V1 to V2 in this subject could possibly be related to speech rate. Subject S3, according to my auditory impression, had the fastest speech rate of all subjects. A faster speech rate would induce more coarticulation in this subject than in the other two subjects, hence less discontinuity in tongue movement throughout VCV sequences would be found. Another reason that would arguably explain this behaviour may be that S3 had lived in the UK for approximately six months by the time of the recording, while the other two subjects had spent a considerably longer amount of time abroad by that time. Another noticeable difference between the subjects in the tongue contour sequence pattern distribution was that subject S1 produced a significant number of troughs in /ip^{j}i/ sequences, while the other subjects did not. In this case, again, speech rate seems to possibly have contributed to this pattern. This subject had a considerably slower speech rate than the other subjects, according to the auditory impression. There are other explanations of the differences described in this paragraph. It may be that the subjects were choosing their own strategies for producing VCV sequences, not necessarily related to speech rate. It seems reasonable to assume that, as long as the target sequences were correctly perceived, the subjects could “afford” some variations in tongue behaviour, including those described above.

All these inter-subject differences considered, a question may arise about the reliability of data gathered from only three participants. Had there been more participants in the study, there could have been more confidence in interpreting statistical results, and less need to address the issue of individual variability. However, the two important patterns that have been observed in this experiment – namely, continuous tongue movement in VCVs with high vowels and tongue raising in /aCa/ sequences – are strongly significant in the results for the three subjects pooled. These two patterns are interesting because they are functional, in that they reflect phonetic characteristics of Russian palatalised and non-palatalised consonants, respectively.

4.4.9. Methodological issues: identifying tongue contour sequence patterns based on the highest point of a curve

It was noted in Section 4.3.7 that some cross-subject differences in whole tongue curves depended on the transducer orientation under the chin, notably on the angle between the transducer and the chin. As one of the methods used in this experiment involved defining tongue contour sequence patterns based on the highest point in the C curve, the issue of transducer orientation needs to be discussed in relation to this method.

As described in Section 3.4.1, every effort was made to make sure that the probe angle was at approximately a right angle to the line of the jaw. Therefore, the probe was expected to be approximately orthogonal to the tongue surface. However, there was some variation in the transducer orientation in relation to the tongue surface. This variation was due to inter-subject differences in the length of the chin and the maximal angle between the chin and the chest. In order to fix the transducer in a comfortable position for the subject, all these parameters had to be taken into account. When the transducer was fixed under the chin, the shadow of the hyoid bone and the shadow of the chin on the ultrasound image were used as reference points, for judging whether the transducer was orthogonal to the line of the jaw. The symmetrical location of the two shadows was considered to represent the situation with a right angle between the transducer and the chin (see Figure 4-1 for an example).

The implications of this methodological issue for the calculations based on the highest point in the tongue curve are as follows. The location of the shadow of the hyoid bone and the shadow of the chin should be taken into account when interpreting the results. Also, all the results of these calculations should necessarily be interpreted together with qualitative analysis of whole curves. In this experiment, the shadows of the hyoid bone and of the chin were close to being symmetrical in two subjects (S1 and S2). This implies that the angle between the transducer and the chin was close to orthogonal. In subject S3, the transducer was slightly rotated clockwise, as evidenced by the position of the two shadows. Midsagittal ultrasound scans of the three subjects in this experiment are presented in Appendix V-1.

The results of the calculations based on the highest point were compared with individual three-spline tokens (15 tokens in each VCV type), and with the average graphs presented in Figures 4-6 – 4-8. This comparison suggests that the results would not have been different, should the probe have been slightly rotated. The only exception is the difference between S3 and the other two subjects. A slight rotation of the transducer clockwise could have altered the distribution of tongue contour sequence patterns in /aba/ and /apa/ sequences in this subject. As noted in Section 4.3.4, in this subject, unlike the other two subjects, the rate of occurrences of antitroughs was above chance in both /aba/ and /apa/ sequences. However, the difference between /aba/ and /apa/ was very small. Examination of three-spline patterns of individual tokens in this subject suggests that if the transducer had been at a right angle to the line of the jaw, this could have resulted in an increased number of “continuous up” tongue movements and a decreased number of antitroughs both in /aba/ and /apa/. This could have brought about the pattern present in the other two subjects: a significant number of antitroughs in /aba/, but not in /apa/.

4.4.10. Relative importance of parts of the tongue in describing tongue movements

Throughout this chapter, movements of specific parts of the tongue have been mentioned: for example, dorsum, root, blade, “back third”, “front third”, etc. Also, movements of the tongue in general have been referred to: “tongue lowering”, “tongue raising”, etc. In this section, based on the results of this experiment, possible specification of parts of the tongue is discussed: whether parts of the tongue may be specified for particular positions, whether they can move independently, and how they are related to “the whole tongue”.

In Section 4.4.8, a difference between the whole tongue contours in the subjects S2 and S3 in /iC^{j}i/ sequences was observed. This difference can be described using the terms introduced in Iskarous (2005). These terms allow us to describe the movement of the whole tongue contour by coordinated movements of its parts. In Figure 4-6d, in S2, there occurs a “pivot” pattern of tongue movement, with a pivot point at the middle of the tongue contour. The pivot pattern is formed by the front half of the tongue continuously moving upwards (with a very small discontinuity: the very front part of the blade in /ip^{j}i/ being lower in the C curve than in the vowel curves), and the back half of the tongue continuously moving forwards. In S3, there is an “arch” pattern of tongue movement, with no such a pivot point. The arch pattern is reached by a continuous raising of all individual parts of the tongue (again, with a very small discontinuity: the very front part of the blade in /ip^{j}i/ being lower in the C curve than in both vowel curves). The difference between these two patterns is especially clear in the /ib^{j}i/ sequence in these two subjects; Figures 4-6c and 4-6e, respectively.

In the case of /uCu/ sequences (Figure 4-7), the term “pivot” mentioned above is also quite convenient for describing the general tongue movement. The most clear example is /ubu/ in S1 (Figure 4-7a): the tongue is moving continuously throughout the

VCV, with a pivot point in the middle; the back half of the tongue is continuously raising, while the front half of the tongue is moving backwards. In other words, we can say that throughout the /ubu/ sequence, the tongue is continuously moving backwards and upwards. This pattern is quite typical, in both /upu/ and /ubu/, and it is followed to a different degree in all three subjects.

Another reason for describing parts of the tongue separately is the method of analysing VCV coarticulation used in this experiment. The fact that the tongue contour sequence patterns are identified based on the highest point in the C curve, makes it easy to describe the trough pattern as “lowering”, and the antitrough pattern as “rising”. So the measurement procedure is a strong factor inducing the use of word combinations like “tongue raising”. However, there is a complicated movement of the tongue, and only the dorsum is definitely raising between the two vowels when antitroughs are observed. The back and the front thirds of the tongue contour are not displaced in the same direction as the middle third of the tongue.

At this stage of the work, the following observation can be made about displacement of parts of the tongue, in relation to the whole tongue contour. The data analysed in this experiment suggest that when the whole tongue is described as being “further forward” or “further backward”, there occurs displacement of the front and back parts of the tongue (normally called in this work “blade” and “root”, respectively) along the x axis. And when the whole tongue is described as being “higher” or “lower”, there occurs displacement of the middle part of the tongue (normally called “dorsum” in this work) along the y axis. So analysing the displacement of the back third, the middle third and the front third of the tongue contour separately allows for describing linguistically relevant tongue behaviour. A conclusion that we can make so far is that describing displacement of these three tongue parts separately is functionally very important.

In Section 5.4.9, the question of specification of tongue parts will be pursued, with more evidence coming from the data analysed in Experiment 2.

4.5. Summary

Experiment 1: Russian bilabial stops

This experiment was aimed at studying coarticulation in Russian VCV sequences with bilabial stops, and the vowels /i/, /u/, /a/. Several hypotheses, based on the existing literature on VCV sequences with bilabial consonants, were formulated and tested. The notion of CR was used in interpreting the results.

It was shown that the tongue shape during a bilabial stop closure in Russian VCV sequences varies greatly, according to the tongue shape for the surrounding vowels. Within the framework of the DAC model, these findings were interpreted by claiming that the degree of resistance to vocalic coarticulation in bilabial stop consonants in Russian VCV sequences was lower than maximal.

Continuous tongue movement from V1 to V2 was observed in /iC^{j}i/ and /uCu/ sequences, and tongue raising was observed between the two vowels in /aCa/ sequences. These results were interpreted as a manifestation of CR by the consonants, motivated by the phonologically based requirements on tongue position for the consonant production.

Stress influence on the tongue shape during the production of the vowels was demonstrated. This result was described in terms of the DAC model by saying that the DAC values of the vowels depend on whether the vowels are stressed or unstressed: the DAC value of the unstressed vowel is lower than the DAC value of the stressed vowel.

Syllable boundary influence on coarticulation in V#CV sequences was shown. Namely, the C curve was more similar to the V2 curve than to the V1 curve. Within the DAC model, these results were interpreted as evidence that the DAC value of the intervocalic consonant in V#CV sequences was influenced by the syllable affiliation of the consonant: the DAC value of the consonant was smaller in relation to V2 than in relation to V1.

In this experiment, quantitative analysis was used for calculating and comparing the numbers of different tongue contour sequence patterns, based on a single point in one curve. Comparison of whole tongue contours was done qualitatively. The results of this qualitative analysis made it possible to formulate the tasks that needed to be solved with quantitative methods in the future. Quantitative Matlab-based methods for measuring and comparing whole tongue curves were designed after analysing the data in this experiment. In Chapter 5 and Chapter 6, some issues raised in this chapter are pursued, and the quantitative procedures designed on the basis of this experiment's findings and observations are applied to analysis of the whole curve.

5. EXPERIMENT TWO:
BRITISH ENGLISH VhV SEQUENCES

5.1. Introduction

This experiment was aimed at measuring resistance to lingual coarticulation in a consonant that has even fewer specifications on its production than bilabial consonants. The consonant [h] "usually denotes a voiceless transition into (or, in some languages, out of) a syllable. Its place of articulation depends on the adjacent sounds" (Ladefoged 2001, p. 254). In this experiment, VhV sequences with the vowels /i/, /u/ and /a/ in British English are studied, and the observed coarticulatory patterns in VhV sequences are compared with the results of the first experiment with bilabial consonants.

The first experiment demonstrated strong coarticulation in bilabial consonants with surrounding vowels. One of the aims of this experiment was to demonstrate, using ultrasound data from British English, that the tongue shape during /h/ in symmetrical VhV sequences varies according to the tongue shape for the surrounding vowels.

It was shown in Experiment 1 that in Russian VCV sequences with bilabial consonants, there was a lack of discontinuity in coarticulation in /iC^{j}i/ and /uCu/ sequences, as related to the results of previous research using data from other languages. The pattern found in our data may have been brought about by the tongue position specification due to phonological palatalisation and phonetic velarisation and pharyngealisation of Russian consonants. In /h/ in English, there are arguably no phonological or phonetic constraints on tongue position. So we could expect that no discontinuity occurs between two identical high vowels in a VhV sequence.

In the first experiment, discontinuity in coarticulation was found in /aCa/ sequences, and it is possible that it occurred because of the phonetic velarisation and pharyngealisation of the intervocalic consonant. Another reason contributing to this pattern may have been jaw raising to assist lip closure for the consonant occlusion causing the tongue to raise between the vowels. In /h/, there are arguably no requirements on tongue or jaw position, so no tongue raising or lowering should be expected between two identical low vowels in a VhV sequence.

While there are no requirements on the supralaryngeal articulators during the production of /h/, differing from those for the surrounding vowels, there may possibly be some requirements on these articulators, coming from the suprasegmental structure of speech. Thus, we could expect that in V#hV sequences, the syllable boundary would affect the tongue position for the consonant, in relation to the two vowels. In Experiment 1, with bilabial consonants, we qualitatively observed a tendency for the consonant curve to be closer to the second vowel's curve than to the first vowel's curve. One of the aims of this experiment was to find out whether a similar tendency is observed in VhV sequences.

In this experiment, the analysis of the ultrasound tongue curves was more technically advanced than in the previous experiment, in that quantitative analysis of tongue displacement along a vertical measure bar was conducted, and Matlab-based quantitative methods of whole curve analysis, designed as part of this experiment, were applied for comparing sets of curves.

5.1.1 Hypotheses

The experiment was designed to test the following hypotheses:

1. **Tongue curves for /h/ in different vowel contexts will be significantly different from each other.** If the hypothesis is supported it will be concluded that /h/ is coarticulated with neighbouring vowels, and that its resistance to lingual coarticulation is lower than maximal. The hypothesis will be refuted if tongue curves for /h/ in different vowel contexts are not significantly different from each other. If this result occurs, it will be concluded that /h/ is not coarticulated with neighbouring vowels in VCV sequences, and that its resistance to lingual coarticulation is absolute.

2. **There will be a continuous tongue movement between the two vowels in /ihi/, /uhu/ and /aha/[5] sequences: specifically, the highest point of the tongue in the C curve will be between the two vowel curves.** If the hypothesis is supported it will be concluded that there is no discontinuity in coarticulation in British English /ihi/, /uhu/ and /aha/ sequences. The hypothesis will be refuted if the highest point in the C curve is consistently higher or lower than the two vowel curves. If this result occurs, it will be concluded that there occurs a discontinuity in coarticulation in British English /ihi/, /uhu/ and /aha/ sequences.

3. **In a V#CV sequence, there will be a significant difference between the V1-C distance and the C-V2 distance: specifically, the distance between the V1 curve and the C curve will be significantly greater than the distance between the C curve and the V2 curve.** If this hypothesis is supported, it will be concluded that there is a syllable boundary influence on the consonant resistance to the surrounding vowels, and that the consonant is less resistant to V2, the vowel belonging to the same syllable of the V#CV, than to V1. This hypothesis will be refuted if there is no significant difference between the V1-C distance and the C-V2 distance, or if the V1-C distance is significantly smaller than the C-V2 distance. These outcomes would mean, respectively, that the consonant's resistance to the surrounding vowels is not influenced by the syllable boundary in V#CV sequences, or that the consonant is more resistant to V2 than to V1.

5.2. Method

5.2.1 Experimental items

The experimental data were two-syllable nonsense sequences /ihi/, /uhu/, /aha/, in the carrier phrase "I said … too". The stimuli were presented to the subjects in English

---

[5] Note that the symbol "a" is used in this chapter to represent the Southern British English low back vowel /ɑ/ and the Standard Scottish English low front vowel /a/.

orthography, in the following way: “eehee” for /ihi/, “oohoo” for /uhu/, and “aha” for /aha/. Stress was not marked.

5.2.2 Subjects

The subjects were three native speakers of British English, two women and one man. The male subject (S2) and one female subject (S3) were speakers of Southern British English. Subject S1 was a speaker of Standard Scottish English. There was no special reason for recording speakers with different accents of English. The accent was not expected to influence the results: irrespective of different vowel qualities in these varieties of English, continuous tongue movement throughout VhV sequences (i.e., absence of troughs) was expected to occur in all the subjects.

5.2.3 Instrumentation and recording procedure

The details of the recording procedure for this experiment are as described in Section 3.4.1, together with the general description of the other two experiments’ recording setup.

There were 15 tokens in each stimulus type. The total number of VCV sequences recorded and analysed in this experiment was 135 tokens.

The order of presentation was the following. One repetition of each of the three sentences was collected as a block, before moving on to the second block, and so on. The sentences in each block were presented in the same order. The order was: “I said eehee too”; “I said oohoo too”; “I said aha too”.

The participants were given a printout of the sentences, for some pre-recording practice, as described in Section 3.4.1 for all three experiments. The subjects were asked to produce the sentences at a comfortable speaking rate. They were instructed to produce the two syllables of the VCV sequence with an equal stress.

5.2.4 Ultrasound software analysis, annotations and splines

Annotating the waveform and creating splines within the ultrasound analysis software (Articulate Assistant), common to all three experiments in this work, was described in

Section 3.4.4. In this experiment, the V1 spline was placed at the mid-point of V1. The C spline was placed at the mid-point of the consonant /h/. The following criteria were used for defining the duration of the consonant /h/. The offset of periodicity for V1 and the onset of periodicity for V2 were considered to be the consonant onset and offset, respectively. The V2 spline was created at the V2 annotation point described in Section 3.4.4. An illustration of the annotations and spline drawing in this experiment is given in Figure 5-1.

Figure 5-1. Illustration of the three annotation points and a spline drawn at the tongue contour, corresponding to the annotation at the mid-point of the consonant /h/.

After annotating the waveform and drawing the splines, the xy data for each curve were exported from Articulate Assistant into a text file, and then imported into Matlab for plotting (see Section 3.4.4.2 on the details of importing tongue curves into Matlab).

Non-Matlab-based quantitative analysis conducted in this experiment will be described in Sections 5.2.7 and 5.2.8. One procedure was similar to the one used in the first experiment (Chapter 4), and it consisted in analysing the distribution of tongue contour sequence patterns. Another procedure consisted in calculating tongue displacement based on a vertical measure bar imposed on the curves.

Matlab-based analysis of whole curves was introduced in this experiment. The text files with xy data for each curve exported from Articulate Assistant were imported into Matlab for comparison of different sets of curves. Calculations in Matlab were done for each speaker separately. Matlab calculations specific to this experiment are described in detail in Sections 5.2.6 and 5.2.9. For the key Matlab procedures described earlier, in Chapter 3, references to relevant sections of the thesis will be provided below where necessary.

5.2.5 Qualitative observation of whole tongue contours

Some qualitative analysis of the curves was made in this experiment. Plots representing average tongue curves were used for portraying whole tongue contours. Each of these plots contained average tongue curves for V1, C and V2 belonging to the same VCV sequence (see Section 3.4.5 for the details of the averaging procedure and for an example of an average curve).

5.2.6 V-on-/h/ coarticulation, analysis in Matlab

The procedure of comparing sets of tongue curves for significant differences was designed, in order to measure V-on-/h/ coarticulation. V-on-/h/ coarticulation was measured by comparing ultrasound curves for the consonant /h/, in the context of the three vowels, /i/, /u/ and /a/. The sets of 15 repetitions of /h/ were compared in the following pairs of vowel environments: /i/ versus /u/, /i/ versus /a/, and /u/ versus /a/. The sets of curves were compared for significant difference, using the procedure based on the Nearest Neighbour technique (see Section 3.4.6 for details of the technique). The procedure described in Section 3.4.7.2 was used: comparing one set of across-group distances to two sets of within-group distances. A Univariate ANOVA was conducted in

SPSS for each pair of vowels separately. The Games-Howell Post Hoc test was used to check for significant differences. If the test showed significant differences between across-group variation and both within-group variations, at the 0.05 level, then the distance between the two sets of curves was considered significant, and it was concluded that there was a V-on-/h/ effect.

5.2.7 Quantitative distribution of tongue contour sequence patterns

The distribution of different tongue contour sequence patterns was calculated for VhV sequences in the three subjects. The criteria for identifying the patterns were the same as used in Experiment 1 (see Section 4.2.6 for the description). The statistical procedure employed to test the numbers of occurrences for significance was the same as the one used in Experiment 1 (see Section 4.2.7 for a short description, and Appendix IV for details). The numbers of troughs, antitroughs and continuous tongue contour sequence patterns (see Section 4.2.6 for definition of these terms) were calculated and compared across subjects and across different vowel environments.

5.2.8 Measurement of tongue displacement along a vertical bar

The extent of tongue displacement in the four different tongue contour sequence patterns was calculated and compared across subjects and across different vowel environments.

The extent of displacement was measured using the following procedure, carried out on the original curves, in Articulate Assistant. The highest point of the tongue contour was identified in the consonant curve, and then compared along a vertical line (called “measure bar”) with the points at which the vertical line intersected the V1 curve and the V2 curve. The extent of displacement from V1 to C was defined as the difference between the V1 curve and the C curve along the measure bar. The extent of displacement from C to V2 was defined as the difference between the C curve and the V2 curve along the measure bar. The measurement procedure is illustrated in Figure 5-2.

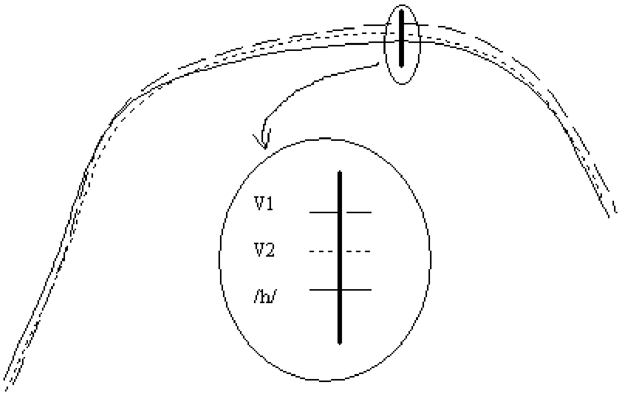


Figure 5-2. Measurement of distances along the vertical measure bar.

The graphs used to represent distances of tongue movement were produced in Matlab. An example of the graph is presented in Figure 5-3, and explained below. The y axis represents tongue displacement in millimetres, starting from V1. The first (left) arrow indicates mean tongue displacement from V1 to /h/ over several repetitions (called "first displacement" in Sections 5.3.5 and 5.3.6). The extent of this displacement is indicated by the distance along the y axis from "V1" to "h". Error bars around "h" represent one standard deviation in the first displacement. The second (right) arrow shows mean tongue displacement from /h/ to V2 (called "second displacement" in Sections 5.3.5 and 5.3.6). The distance along the y axis from "h" to V2 indicates the extent of the displacement, and error bars around "V2" represent one standard deviation in the second displacement. The consonant is taken as a zero on the y axis for convenience of reading the graph.

A Univariate ANOVA was conducted in SPSS to look for significant differences in tongue displacement between vowel environments and between subjects. The Tukey HSD Post Hoc test was used to show where exactly these differences occurred.

Experiment 2: British English /h/

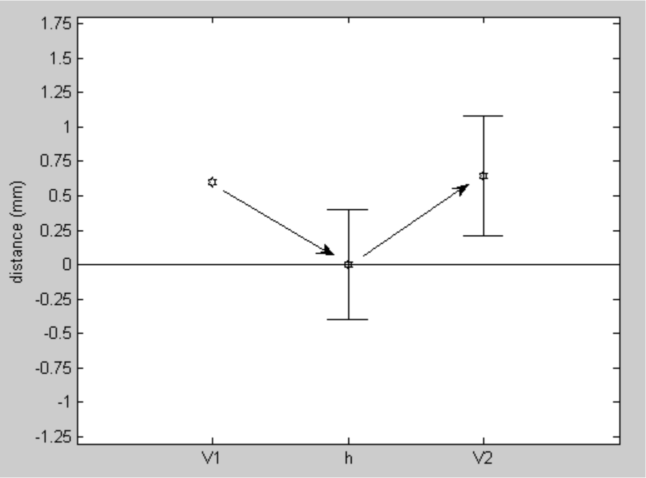


Figure 5-3. Representing distances of tongue movement along the vertical measure bar.

5.2.9 Comparison of V1-C and C-V2 distances, analysis in Matlab

The procedure for comparing distances between one pair of curve sets and another pair of curve sets was designed in this experiment, in order to compare the extent of two different tongue displacements for significant differences, based on whole curves. For calculating the difference between V1-C and C-V2 distances, the procedure described in detail in Section 3.4.7.1 was used. Average nearest neighbour distances were calculated, for /ihi/, /uhu/ and /aha/ sequences, between V1 and C, and between C and V2. The V1-C distance was indicated by the average distance between the set of 15 repetitions of V1 and the set of 15 repetitions of C. The C-V2 distance was indicated by the average distance between the set of 15 repetitions of C and the set of 15 repetitions of V2.

Statistical comparison was made in SPSS. A Univariate ANOVA was conducted separately for each subject. Average nearest neighbour distances were compared for V1-C versus C-V2 (the independent variable was called syllable affiliation of the vowel), and for the three VCV types, /ihi/, /uhu/ and /aha/ (the independent variable was called VCV type). Pairwise comparison with the Bonferroni adjustment was used to look for significant differences depending on the syllable affiliation of the vowel. If the C-V2 distance was significantly smaller than the V1-C distance, at the 0.05 level, it was concluded that there was a syllable boundary effect on the coarticulation in the VCV sequence. The Games-Howell Post Hoc test was used to check for significant differences depending on the VCV type. If differences were significant, at the 0.05 level, it was concluded that V1-C and C-V2 distances in VhV sequences depended on the vowel environment.

A separate Univariate ANOVA was run, to find out whether there were differences between the subjects. Pairwise comparison, with the Bonferroni adjustment, was used to look for significant differences. If differences occurred, at the 0.05 level, it was concluded that V1-C and C-V2 distances in VhV sequences depended on the subject.

5.3. Results

5.3.1 V-on-C coarticulation in /h/

In Figure 5-4, tongue curves for 15 repetitions of /h/ in the context of the three vowels are presented, for subject S2. This figure illustrates the pattern that was observed in all three subjects.

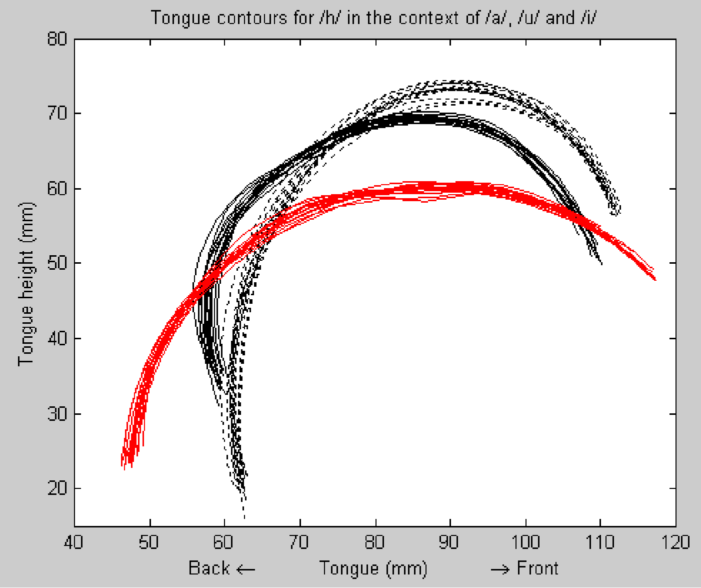
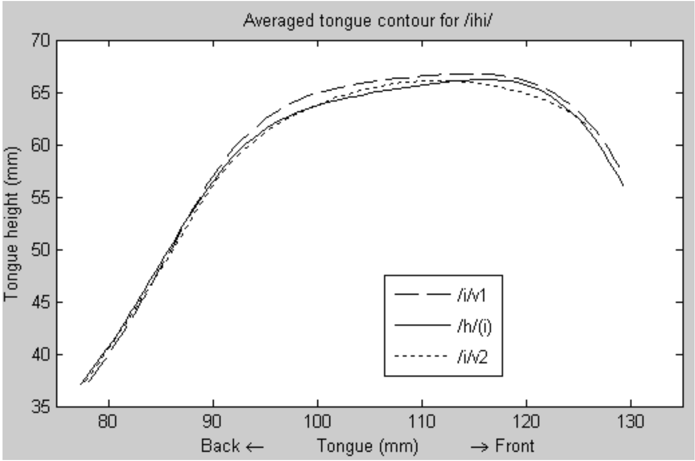


Figure 5-4. Tongue contours for /h/ in the context of three different vowels, subject S2. Red lines – fifteen tokens of /h/ in the context of /a/; solid black lines – fifteen tokens of /h/ in the context of /u/; dashed black lines – fifteen tokens of /h/ in the context of /i/.

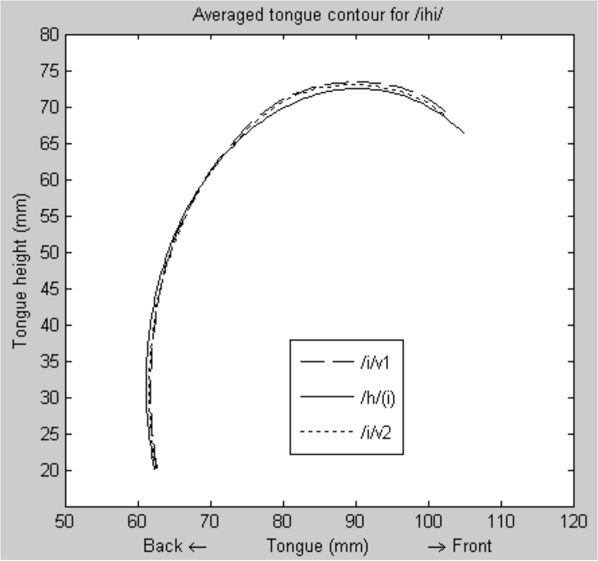
The sets of curves were compared for two subjects, S2 and S3, for each pair of vowel environments: /i/-/u/, /i/-/a/, and /u/-/a/. In the event, it was only possible to do quantification for these two subjects, because during the recording of subject S1, there occurred a slight shift forwards of the helmet with the transducer. The results of the Univariate ANOVA for all three pairs of vowels show a significant effect for both subjects analysed. In S2, for /i/-/u/, $F = 3909.48$, $df = 2$, $p \leq 0.001$; for /i/-/a/, $F = 25862.92$, $df = 2$, $p \leq 0.001$; for /u/-/a/, $F = 15390.36$, $df = 2$, $p \leq 0.001$. In S3, for /i/-/u/, $F = 1770.50$, $df = 2$, $p \leq 0.001$; for /i/-/a/, $F = 27294.09$, $df = 2$, $p \leq 0.001$; for /u/-/a/, $F = 11961.92$, $df = 2$, $p \leq 0.001$. The Games-Howell Post Hoc tests demonstrate that across-group distances are always significantly greater than within-group distances, at the 0.05 level. This means that the /h/ curves are significantly different from each other in different vowel environments.

Below, in Figures 5-5 – 5-7, the curves for V1, /h/ and V2 belonging to the same VCV sequence are displayed on the same graph. These graphs are displayed here for making some visual observations about coarticulatory patterns in VhV sequences, based on whole tongue curves. Each of the three curves on each graph is an average curve over 15 tokens. A separate graph is made for each VhV type, for each subject individually.

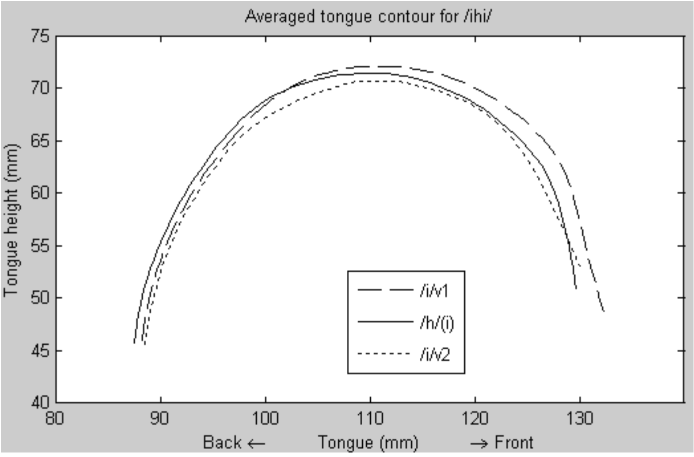
In Figure 5-5, average V1, C and V2 curves for /ihi/ sequences are displayed.



a) S1, /ihi/



b) S2, /ihi/



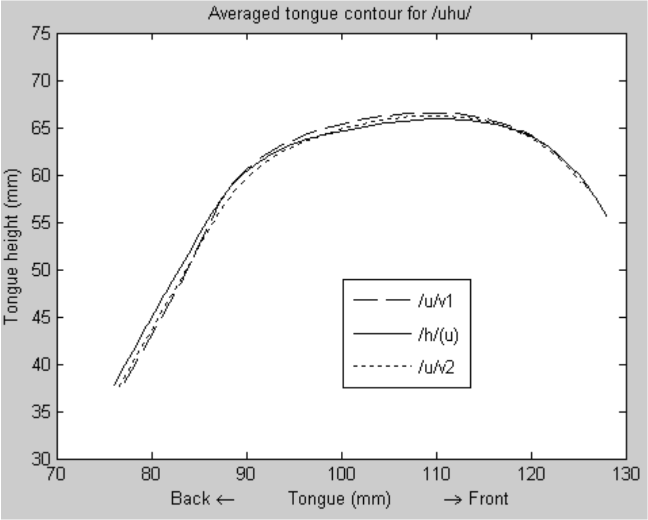
c) S3, /ihi/

Figure 5-5. Average V1, C and V2 curves for /ihi/ sequences, for the three subjects: a) S1; b) S2; c) S3. Dashed line – V1 curve; solid line – C curve; dotted line – V2 curve.

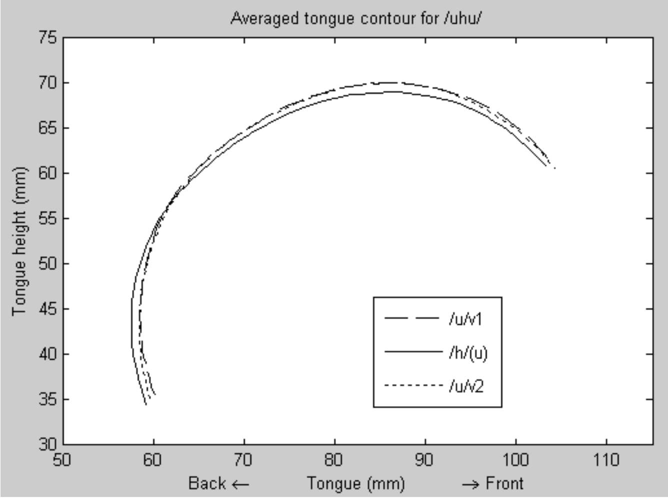
In Figure 5-5, the contours for V1, /h/ and V2 are very close together. This fact confirms the results presented in Figure 5-4, that the tongue shape for the consonant in a VCV sequence varies according to the surrounding vowels' tongue shapes.

Average /uhu/ tongue contours for the three subjects are presented in Figure 5-6.

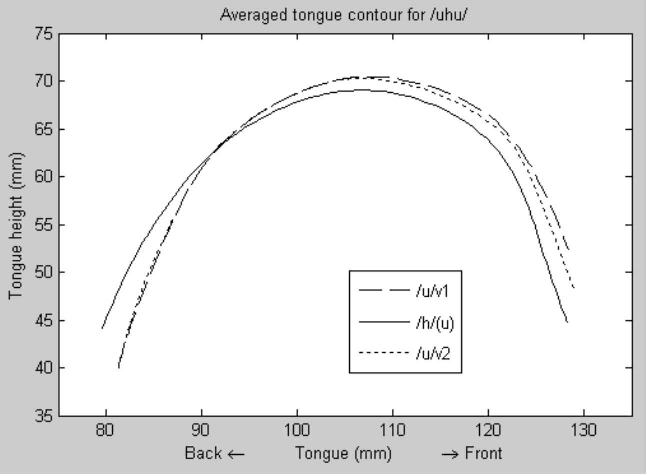
Experiment 2: British English /h/



a) S1, /uhu/



b) S2, /uhu/

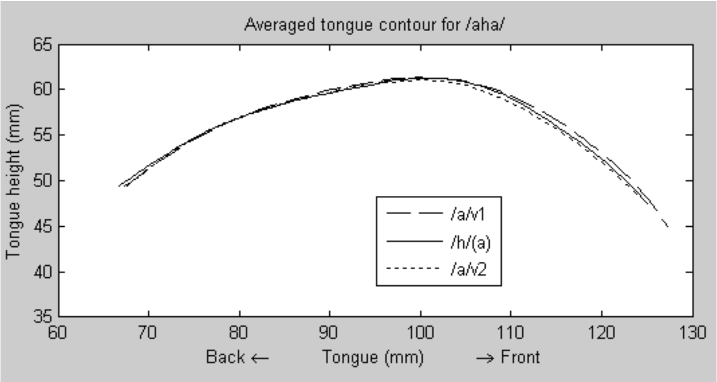


c) S3, /uhu/

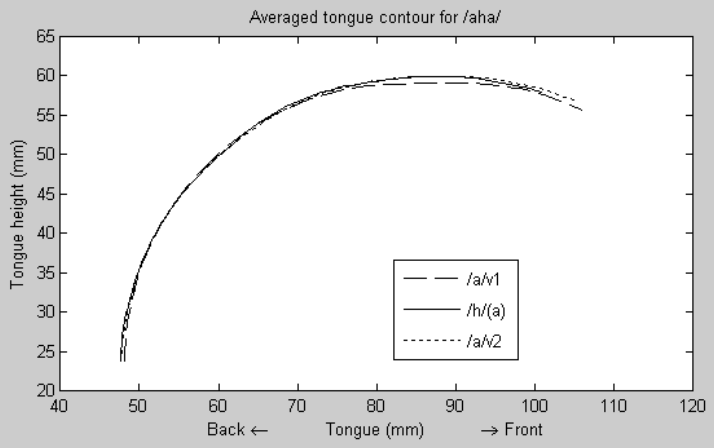
Figure 5-6. Average V1, C and V2 curves for /uhu/ sequences, for the three subjects: a) S1; b) S2; c) S3. Dashed line – V1 curve; solid line – C curve; dotted line – V2 curve.

Again, as in /ihi/ sequences, the three curves in the VCV are rather close together. In S3, both in /ihi/ and in /uhu/ sequences, the consonant curve appears to be the furthest away from the two vowel curves, as compared with the other subjects. We will come back to this cross-subject difference in Section 5.3.7.

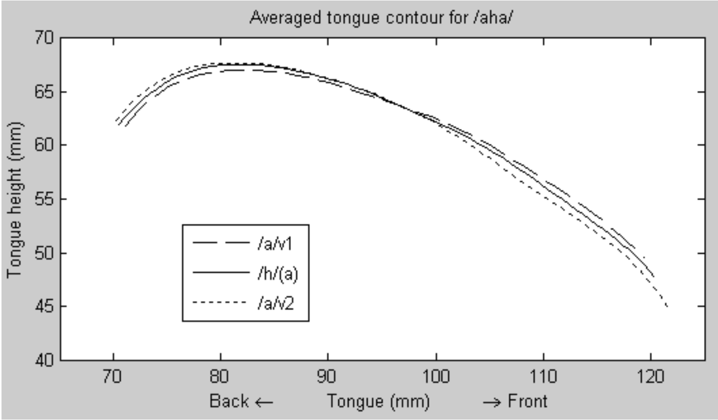
Average /aha/ tongue contours for the three subjects are presented in Figure 5-7.



a) S1, /aha/



b) S2, /aha/



c) S3, /aha/

Figure 5-7. Average V1, C and V2 curves for /aha/ sequences, for the three subjects: a) S1; b) S2; c) S3. Dashed line – V1 curve; solid line – C curve; dotted line – V2 curve.

In /aha/ sequences, as well as in the two other vowel contexts, the three tongue curves displayed on each graph are rather close together. Again, as in the other two vowel contexts, S3's consonant curve appears to be further away from the two vowel curves than is the case for the other subjects. This cross-subject difference will be discussed in Section 5.4.7.

The general observation about Figures 5-5 – 5-7 is that the contours of V1, C and V2 are close together in VCV sequences. This observation is consistent with the data presented in Figure 5-4, where the shapes of the C curve were shown to be significantly different, depending on the vowel environment. All these data show that the tongue surface shape during /h/ varies greatly according to the identity of the surrounding vowels.

5.3.2 Qualitative analysis of tongue movements in VhV sequences

It was noted in the previous section that the three tongue curves are very close together in VhV sequences. However, there are some differences between the three curves, reflecting the movement of the tongue from V1 to V2, through the consonant. These differences are described here.

In /ihi/ sequences (Figure 5-5), in all the subjects, the tongue root is further back in the C curve than in both vowel curves. This is evidenced by the fact that the x values of the back part of the C curve are smaller than the x values of the the back part of each vowel curve with the same y values. In S1, the portion of the C curve that is further back than the V curves takes approximately one fourth of the whole tongue contour; in S2 and S3, over a third of the whole contour**.** While the tongue root moves backwards for producing /h/ in-between the two /i/ vowels, the rest of the tongue (mid and front) seems to follow it. In S1 and S3, the front part of /h/ curve is further back than V1. This is demonstrated by the fact that the x values of the front part of the C curve are noticeably smaller than the x values of the the front part of the V1 curve with the same y values. The middle part of the /h/ curve in S1 and S3 is lower than the V1 curve. This is evidenced by the smaller y values of the middle part of the /h/ curve than the y values of the V1 curve with the same x values. In S2, there is a similar tendency to the one described for S1 and S3. In S2, the transducer was rotated anti-clockwise in comparison with the other subjects (see more on the issue of transducer orientation in Section 5.4.8). So the description of tongue movements for this subject in terms of xy coordinates is slightly different from the description for S1 and S3: in S2, the front half of the tongue in /h/ is lowered, in relation to both vowels.

One of the common features in all the subjects for /uhu/ sequences is that the tongue root (approximately one third of the tongue contour in all the graphs in Figure 5-6) is going backwards between the two vowels' steady states. Another common feature across subjects is that the middle part of the tongue lowers for the consonant. A difference between the subjects is that the front part of the tongue in /h/ is below both vowels in S2 and S3, and above both vowels in S1.

In /aha/ sequences (Figure 5-7), in all the subjects, the curves for V1, C and V2 appear to be closer together than in the VCV sequences with high vowels. This difference will be quantified in Section 5.3.7, and discussed in Section 5.4.3. A common feature across all the subjects in /aha/ sequences is that the front part of the tongue moves continuously from V1 to V2. The direction of movement, however, is not the same in all the subjects. In S1 and S3, the tongue blade is continuously lowering, more so in S3. In S2, in contrast, the tongue blade is higher for V2 than for V1. In /aha/ sequences, for all the subjects, there is a noticeable difference between the curves in the back third of the tongue curve. S3 differs from both other subjects in that there is a continuous backward movement from V1 to V2. In S1 and S2, the position of the tongue root is very close to both vowels, and the very back of it is slightly further backwards for /h/ than for both vowels.

These visual observations will be referred to in the following sections, where tongue displacement along a measure bar is analysed. Discussion of these observations in relation to the measure bar calculations is offered in Section 5.4.8.

5.3.3 Distribution of tongue contour sequence patterns

In Table 5-1, the numbers of occurrences of tongue contour sequence patterns are presented according to vowel type, for all the subjects together. For definitions of the patterns, see Section 4.2.6. For details of statistical calculations, see Appendix IV. In each vowel group, there are 45 tokens. According to the binomial experiment conditions, in a case of 45 repetitions, a pattern occurs significantly above chance level if its number of occurrences is equal to or more than 19, and significantly below chance level if its number of occurrences is equal to or fewer than 4, at $p < 0.05$.

	Trough	Antitrough	Continuous up	Continuous down
i	41	0	1	3
u	43	0	1	1
a	11	16	14	4

Table 5-1. Numbers of occurrences of tongue contour sequence patterns by vowel type, for all the subjects together.

Figure 5-8 illustrates the distribution of the different tongue contour sequence patterns according to vowel type, in the three subjects together. The results show that the rates of occurrence of troughs in the two high vowels, /i/ and /u/, are significantly above chance ($p < 0.001$). The rates of occurrence of continuous patterns in /ihi/ and /uhu/ are significantly smaller than chance (in /ihi/, $p < 0.001$ for “continuous up” and $p < 0.002$ for “continuous down”; in /uhu/, $p < 0.001$ for both continuous patterns). As for antitroughs, in /i/ and /u/ environments there are none.

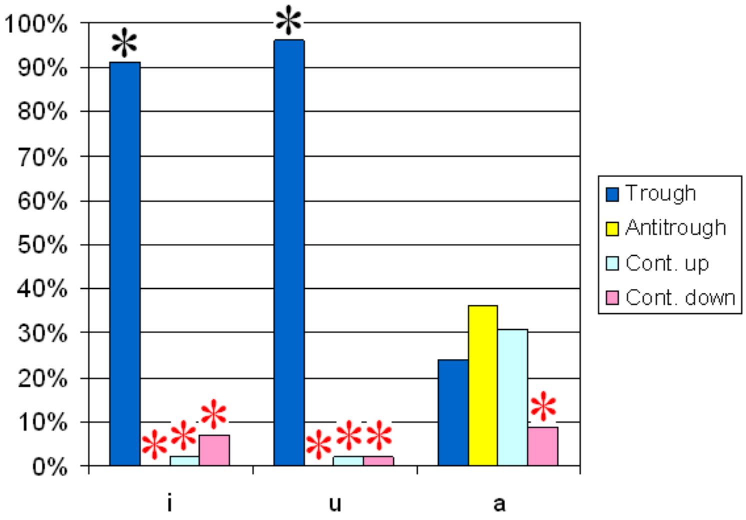


Figure 5-8. Distribution of tongue contour sequence patterns by vowel, for all the subjects together (percentage, out of 45 tokens). A black asterisk above a bar means that the rate of occurrence was significantly above chance. A red asterisk above a bar means that the rate of occurrence was significantly below chance.

In /aha/ sequences, the rate of occurrence of troughs is at chance level, as well as the rate of occurrence of "continuous up" patterns. "Continuous down" patterns are significantly below chance in /aha/ ($p < 0.01$). The rate of occurrence of antitroughs in /aha/ is at chance level.

5.3.4 Distribution of tongue contour sequence patterns, individual results

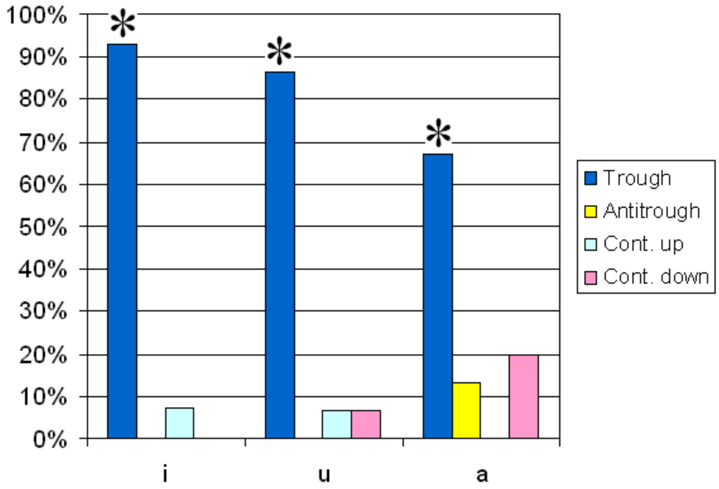
Here, the data are analysed for the three subjects separately. In Table 5-2, the numbers of occurrences of tongue contour sequence patterns are presented by subject and by vowel type. In each group (represented by a row in the table), there are 15 tokens. According to the binomial experiment conditions, in a case of 15 repetitions, a pattern occurs significantly above chance level if its number of occurrences is equal to or more than 9. In this case, no patterns occur significantly below chance, because zero occurrences has a mathematical probability of occurrence of 0.0134, which is above the threshold of 0.0125 resulting from the Bonferroni adjustment (see Section 4.2.7 and Appendix IV for more details).

Experiment 2: British English /h/

		Trough	Antitrough	Continuous up	Continuous down
S1	i	14	0	1	0
	u	13	0	1	1
	a	10	2	0	3
S2	i	14	0	0	1
	u	15	0	0	0
	a	0	8	7	0
S3	i	13	0	0	2
	u	15	0	0	0
	a	1	6	7	1

Table 5-2. Numbers of occurrences of tongue contour sequence patterns, by subject and by vowel type.

The distribution of the four different patterns of tongue movement by subject and by vowel type is presented in Figure 5-9.



a)

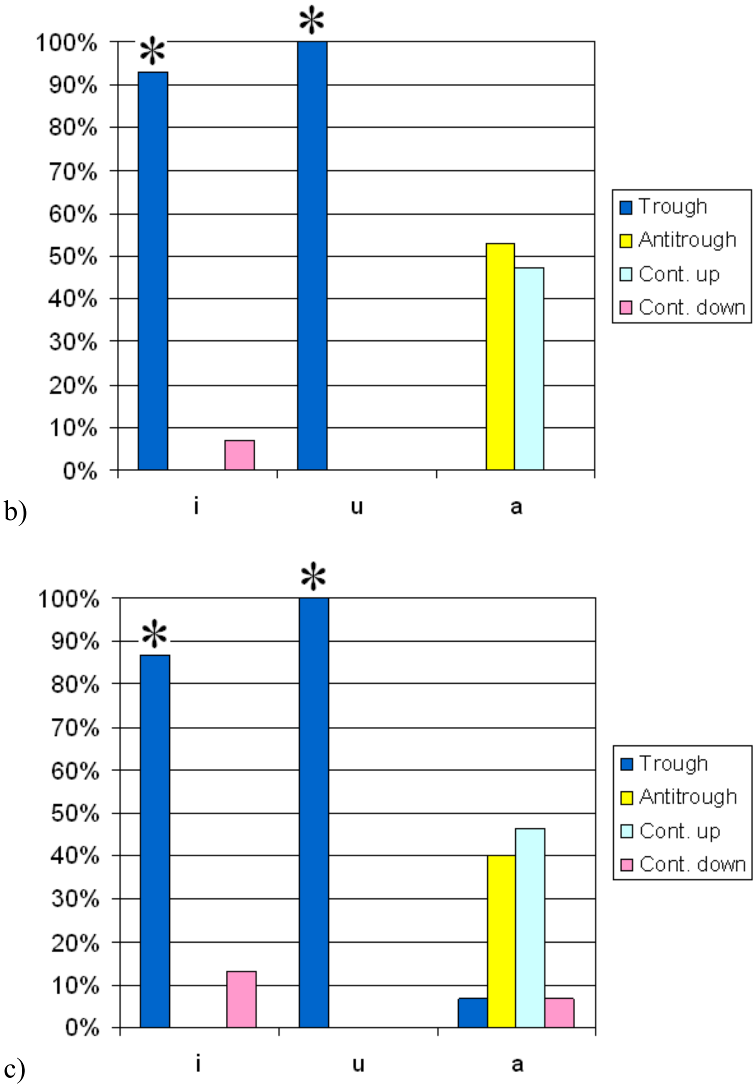


Figure 5-9. Distribution of tongue contour sequence patterns by subject and by vowel type (percentage, out of 15 tokens): a) subject S1; b) subject S2; c) subject S3. A black asterisk above a bar means that the rate of occurrence was significantly above chance.

The most uniform distribution of tongue contour sequence patterns across subjects is found in the /u/ environment. Two subjects, S2 and S3, produce troughs on 100% of all repetitions. The rate of occurrence of troughs in the /u/ environment in S1 is also significantly above chance ($p < 0.001$), but this subject also has some continuous patterns. An example of a trough pattern in /uhu/ from S3 is given in Figure 5-10.

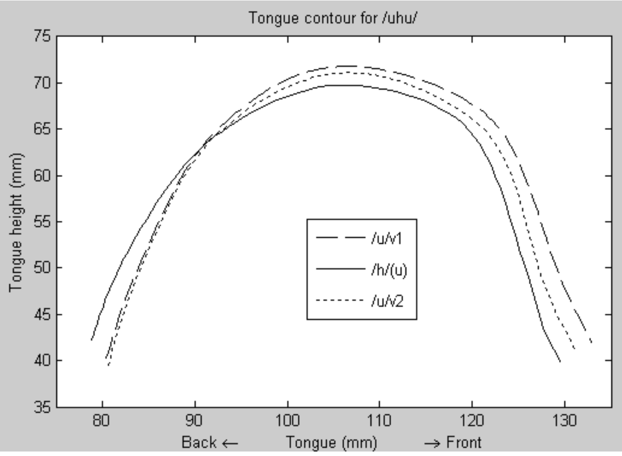


Figure 5-10. An example of a trough in S3, in /uhu/. Dashed line – V1 curve; solid line – C curve; dotted line – V2 curve.

In /ihi/ sequences, none of the subjects has 100% of troughs; continuous patterns are present in all the subjects. Still, the trough pattern occurs at significantly greater than chance level ($p < 0.001$), in all the subjects. An example of a trough pattern in /ihi/ from S1 is given in Figure 5-11.

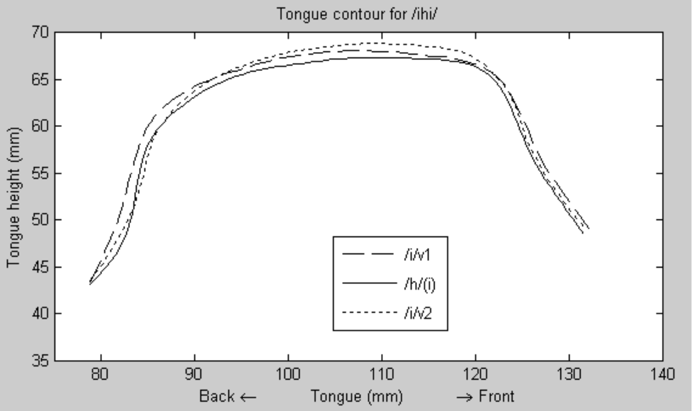
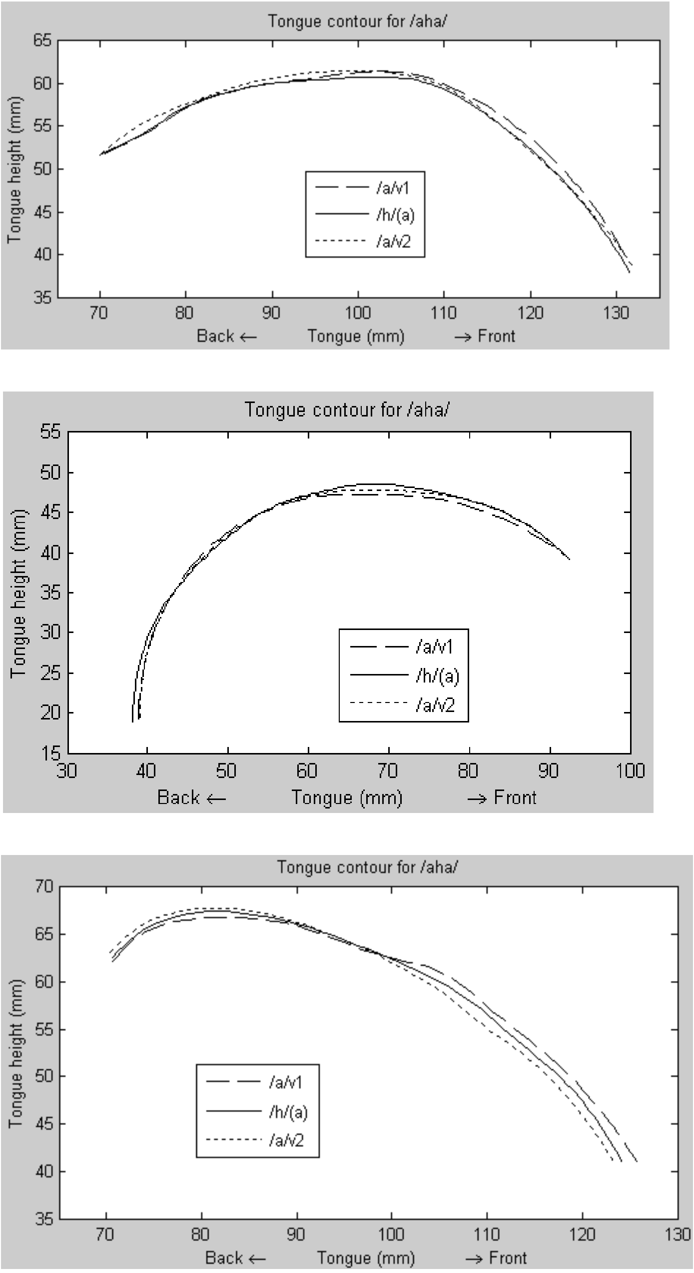


Figure 5-11. An example of a trough in S1, in /ihi/. Dashed line – V1 curve; solid line – C curve; dotted line – V2 curve.

The patterns of tongue behaviour in the /a/ environment differ most across subjects. While S1 has two thirds of the patterns realised as troughs (significantly above chance level, $p < 0.001$), S2 does not produce troughs at all, and S3 only produces one trough out of 15 repetitions. Both S2 and S3 have antitroughs and continuous patterns instead. However, the rates of occurrence of continuous patterns and antitroughs in both subjects are at chance level. In Figure 5-12, there are examples of tongue contour sequence patterns in /aha/ in the three subjects: a trough in S1, an antitrough in S2, and a continuous pattern in S3.



a) S1

b) S2

c) S3

Figure 5-12. Examples of tongue contour sequence patterns in /aha/ for the three subjects: a) S1; b) S2; c) S3. Dashed line – V1 curve; solid line – C curve; dotted line – V2 curve.

Summarising this description, we can say that the three subjects are close in the rate of trough occurrence in /ihi/ and /uhu/: in these two vowel environments, all the subjects have over 80% of troughs ($p < 0.001$ in each subject individually and in all the subjects pooled). In the /a/ contexts, subjects S2 and S3 demonstrate rather similar patterns of tongue behaviour, while S1 is different from both of these subjects. In S1, troughs dominate in /aha/ sequences (67%, which is significantly above chance, $p < 0.001$), while in S2 and S3, antitroughs and continuous upward movement patterns are rather frequent (from 40% to 53%; always at chance level). More on individual differences occurring in the /a/ environment will be presented in Section 5.3.6. The reasons for these individual differences will be discussed in Section 5.4.8.

5.3.5 Tongue displacement along a vertical bar

Extent of tongue displacement was measured, using the procedure described in Section 5.2.8. The results are presented below.

Extent of displacement, averaged over 15 repetitions, is shown in Table 5-3 for all the subjects together.

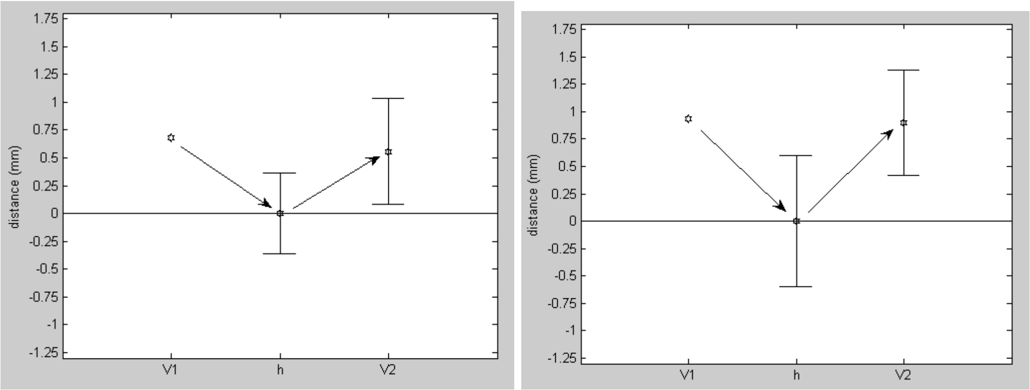
	ihi		uhu		aha	
	first	second	first	second	first	second
S1	-0.68	0.56	-0.94	0.90	-0.38	0.07

Table 5-3. Mean tongue displacements for all the subjects together, in millimetres, from V1 to C (“first”) and from C to V2 (“second”), as calculated along a vertical measure bar. Minus signs in the “first” column indicate that the tongue moved downwards from V1 to C, as measured along the vertical bar.

Note that all tongue movement distances are very small. The extent of tongue displacement ranges from less than 0.1 mm to slightly over 2 mm. Considering these small numbers, some differences may be within the margin of measurement error (cf. Stone 1999, who claims that measurement error on ultrasound images is less than 0.7 mm). By this, we mean that if two independent transcribers create the splines for the same token, the absolute xy values may be different. However, drawing splines is largely based on looking at tongue dynamics along the VCV sequence and capturing the change occurring in time, at the three consecutive time points. So two independent transcribers would be expected to represent the direction of tongue movement in the same way, even though there may be differences in absolute xy coordinates of the splines drawn. Hence, a similar three-contour pattern of tongue displacement (V1-C-V2) would be almost certainly represented by different transcribers. No special investigation of reliability of ultrasound measurements was conducted in this work.

A Univariate ANOVA was conducted, separately for the first displacement and the second displacement, to test whether there was a significant effect of the vowel context and of the individual subjects on the extent of displacement. The effect of the vowel context was significant in both cases (the first displacement: $F = 132.89$, $df = 2$, $p \leq 0.001$; the second displacement: $F = 42.66$, $df = 2$, $p \leq 0.001$).

The distances of tongue movement, for all the subjects together, are shown in Figures 5-13 and 5-14 (for a detailed explanation of the graph, see Section 5.2.8).



a) b)

Figure 5-13. Mean tongue displacement into and out of /h/, for the two high vowels, for all the subjects together: a) vowel /i/ context; b) vowel /u/ context.

Figure 5-13 shows that in the two high vowel environments, the average tongue contour sequence pattern across subjects, based on the tongue displacement calculations, is down (V1-C) followed by up (C-V2), i.e., a trough in tongue movement.

In Figure 5-14, the results of the tongue displacement in /aha/ sequences, for the three subjects pooled, are presented. We see that the average pattern is a continuous tongue movement upwards from V1 to V2. This figure shows that tongue displacements for V1-C and for C-V2 are rather small in /aha/. The Tukey HSD Post Hoc test demonstrates that in /aha/, both displacements are significantly smaller than the V1-C and C-V2 displacements in /i/ and /u/ contexts, at the 0.05 level.

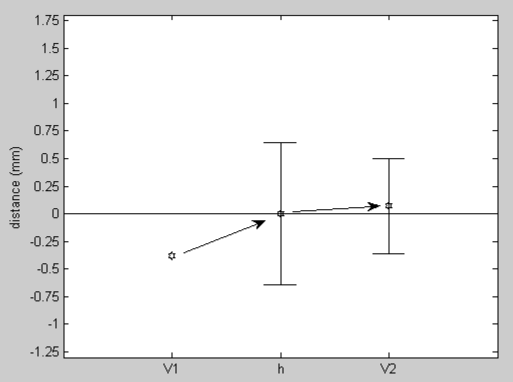


Figure 5-14. Mean tongue displacement into and out of /h/, for all the subjects together, in the vowel /a/ context.

A comparison of these results with the results from Section 5.3.3 is given below. It is interesting to compare the graphs representing distances of tongue movement along a vertical measure bar (Figure 5-13) with the bar plot in Figure 5-8, which represents the distribution of the four different tongue contour sequence patterns. Some tendencies are reflected in both graphs. For example, Figure 5-8 shows that the greatest number of troughs occurred in the /u/ environment, and Figure 5-13 shows the greatest extent of tongue displacement in /uhu/, as compared with the other two vowel environments. The directions of these two displacements in /uhu/ sequences correspond to the definition of trough we gave in Section 4.2.6: first a downward tongue displacement, then an upward tongue displacement. In the /ihi/ context, the tongue also goes down then up (Figure 5-13), but the extent of displacement is smaller than in /uhu/. This is consistent with the distribution of tongue contour sequence patterns: in Figure 5-8, the number of troughs is smaller in the /i/ environments than in the /u/ environments.

In /aha/, the results of the tongue displacement analysis are only partly consistent with the results of the distribution of tongue contour sequence patterns, presented in Section 5.3.3. In Figure 5-8, we see that the most often occurring tongue contour sequence pattern in /aha/ is the antitrough (i.e., the tongue going up then down), though its number of occurrences does not reach significance. Only the second most frequent pattern in Figure 5-8 is "continuous up", the same as the average displacement pattern we see in Figure 5-14.

5.3.6 Tongue displacement along a vertical bar, individual results

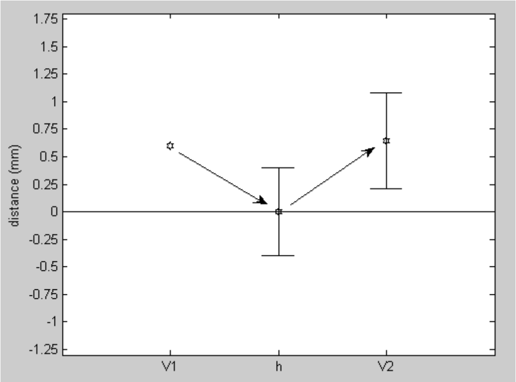
Let us now break these results down by subject. A separate Univariate ANOVA was run for each subject.

The displacements for subject S1 are shown in Table 5-4.

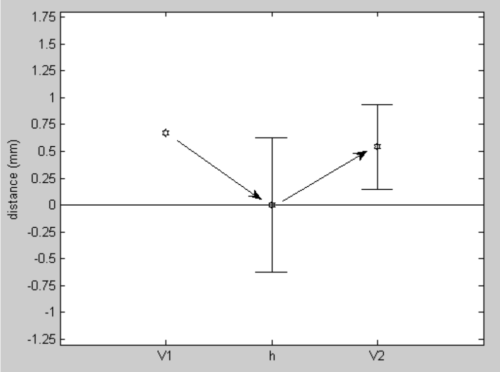
ihi		uhu		aha	
first	second	first	second	first	second
-0.60	0.64	-0.67	0.54	-0.34	0.15

Table 5-4. Mean tongue displacements for subject S1, in millimetres, from V1 to C ("first") and from C to V2 ("second"), as calculated along a vertical measure bar.

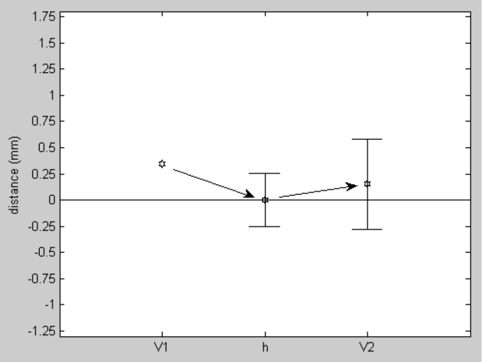
The graphs of tongue displacement in S1 are presented in Figure 5-15.



a) S1, /ihi/



b) S1, /uhu/



c) S1, /aha/

Figure 5-15. Mean tongue displacement into and out of /h/, for the three different vowels, in subject S1: a) vowel /i/ context; b) vowel /u/ context; c) vowel /a/ context.

In this subject, the pattern for the two high vowels is the same as in the across-subject results (Figure 5-13). In /aha/, this subject has a trough tongue movement from V1 to V2. This pattern is not consistent with the pattern for all three subjects pooled, in Figure 5-14.

The results shown in the displacement graphs (Figure 5-15) are quite consistent with the tongue contour sequence pattern distribution in this subject (Section 5.3.4, Figure 5-9a). In Figure 5-9a, in all three vowel contexts, the rate of occurrence of troughs is significantly above chance ($p < 0.001$ for all three vowel environments). The graphs in Figure 5-15 feature a trough movement (down-up) in all three vowel contexts. The extent of displacement in /aha/ is smaller than in the other two vowels (Figure 5-15), and the number of troughs on /aha/ is smaller than in /ihi/ and /uhu/ (Figure 5-9a).

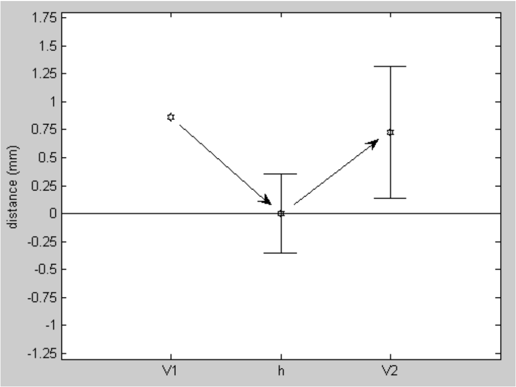
The ANOVA results show that there was no significant effect of the vowel context on the first displacement in this subject, and that for the second displacement, the effect of the vowel context was significant (*F* = 5.71, *df* = 2, $p \leq 0.01$). The results for the first displacement in this subject differ from the results for the three subjects pooled, presented above, in that in S1 there are no significant differences between /a/ and the other two vowels. This is explained by the fact that S1 was producing a lot of troughs in /aha/ sequences, unlike the other subjects. However, /a/ still differs from the other two vowel environments, as evidenced by the second displacement. The Tukey HSD Post Hoc test demonstrates that the extent of the second displacement is significantly different between /a/ and the other two vowel environments, at the 0.05 level, while there are no significant differences between /i/ and /u/.

The displacements for subject S2 are shown in Table 5-5.

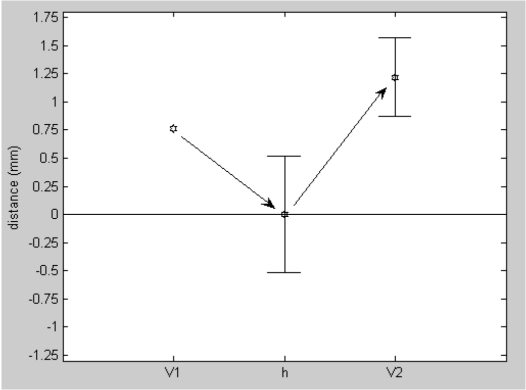
ihi		uhu		aha	
first	second	first	second	first	second
-0.86	0.73	-0.77	1.22	1.02	-0.04

Table 5-5. Mean tongue displacements for subject S2, in millimetres, from V1 to C (“first”) and from C to V2 (“second”), as calculated along a vertical measure bar.

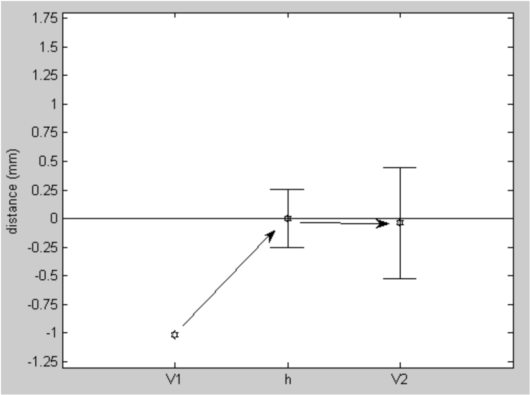
Figure 5-16 shows the results of the tongue displacements for subject S2.



a) S2, /ihi/



b) S2, /uhu/



c) S2, /aha/

Figure 5-16. Mean tongue displacement into and out of /h/, for the three different vowels, in subject S2: a) vowel /i/ context; b) vowel /u/ context; c) vowel /a/ context.

In /ihi/ and /uhu/ in S2, we see the same pattern as in S1. In /aha/, the displacement pattern in S2 is antitrough, i.e., up-down tongue displacement between the two vowels (see an example of the antitrough pattern from this subject in Figure 5-12b). The second displacement in /aha/ in S2 is quite small, and its standard deviation is rather great, so the displacement pattern in /aha/ in this subject is not very different from the pattern for all three subjects pooled (Figure 5-14).

In this subject, like in S1, the results of the tongue displacement measurement are quite consistent with the distribution of tongue contour sequence patterns presented in Section 5.3.4. The rate of trough occurrence in /i/ and /u/ contexts is significantly above chance (Figure 5-9b, $p < 0.001$ for both vowel environments), and the average displacements in Figures 5-16a and 5-16b also have the trough outlook: down (V1-C) – up (C-V2). In /uhu/, this subject has more troughs than in /ihi/ (Figure 5-9b), and the second displacement in /uhu/ is significantly greater than in /ihi/ (at 0.05 level, as evidenced by the Tukey HSD Post Hoc test). In /aha/, the number of antitroughs Figure 5-9b is greater than the number of other patterns; however, the rate of antitrough occurrence does not reach significance. In Figure 5-16c, we see a very small-sized antitrough as the average displacement pattern.

The ANOVA results show that there was a significant effect of the vowel context on both displacements in this subject (first displacement: $F = 110.31$, $df = 2$, $p \leq 0.001$; second displacement: $F = 25.46$, $df = 2$, $p \leq 0.001$).

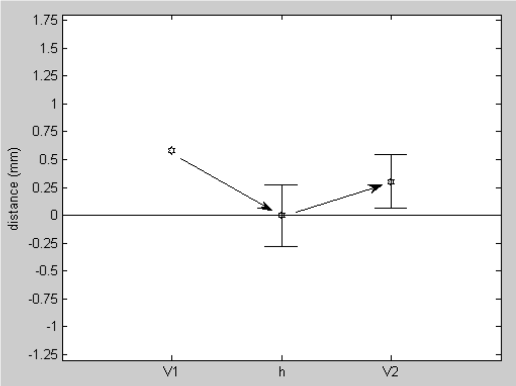
In S2, both the first and the second displacement are significantly smaller in /a/ than in the other two vowel contexts, at 0.05 level, as shown by the Tukey HSD Post Hoc test. This difference between vowel contexts will be referred to again in Section 5.3.7.

The displacements for subject S3 are shown in Table 5-6.

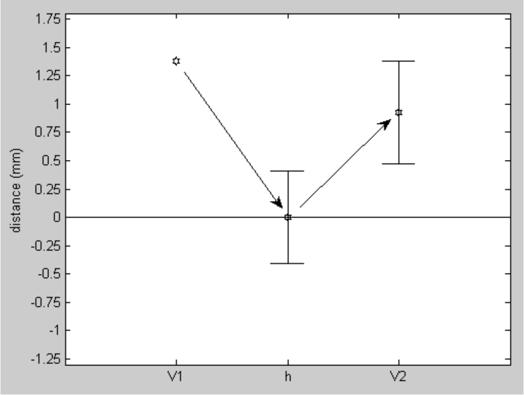
ihi		uhu		aha	
first	second	first	second	first	second
-0.58	0.30	-1.37	0.93	0.45	0.09

Table 5-6. Mean tongue displacements for subject S3, in millimetres, from V1 to C ("first") and from C to V2 ("second"), as calculated along a vertical measure bar.

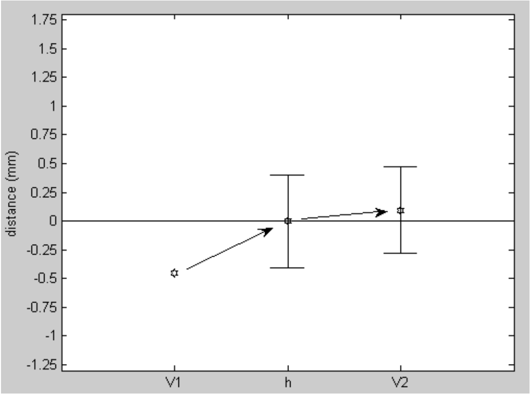
In Figure 5-17, tongue displacements for subject S3 are presented.



a) S3, /ihi/



b) S3, /uhu/



c) S3, /aha/

Figure 5-17. Mean tongue displacement into and out of /h/, for the three different vowels, in subject S3: a) vowel /i/ context; b) vowel /u/ context; c) vowel /a/ context.

In /ihi/ and /uhu/ in S3, just like in the other two subjects, we see the trough pattern. In /aha/, unlike the other two subjects, in S3 we see a continuous tongue movement upwards from V1 to V2.

As well as in the other two subjects, tongue displacements in /ihi/ and /uhu/ in Figure 5-17 are consistent with the pattern presented in Figure 5-9c. In /aha/, this subject behaves much like S2 (compare Figures 5-9b and 5-9c), in that “continuous up” patterns and antitroughs dominate. “Continuous up” tongue contour sequence pattern has the highest number of occurrences in this subject in Figure 5-9c, but neither antitroughs nor “continuous up” patterns reach significance. This distribution is consistent with Figure 5-17, showing a continuous upward tongue displacement in /aha/, with the first movement being greater, and the second movement being smaller.

The ANOVA results show that there was a significant effect of the vowel context on both displacements in this subject (first displacement: $F = 92.98$, $df = 2$, $p \leq 0.001$; second displacement: $F = 21.08$, $df = 2$, $p \leq 0.001$). In S3, all three vowels are significantly different from each other in the first tongue displacement (at the 0.05 level, the Tukey HSD Post Hoc test). In the case of /a/, the mean displacement is upwards, and in the other two vowel contexts the tongue goes downwards from V1 to C, so there is an obvious difference between the two patterns, opposing /a/ to the other two vowels. The second displacement shows significant difference between /u/ and the other two vowels (at the 0.05 level, the Tukey HSD Post Hoc test). The vowels /i/ and /a/ happen to be not significantly different from each other in this case.

5.3.7 Comparison of V1-C and C-V2 distances

A qualitative observation was made in Section 5.3.1, that the V2 curve often appears to be closer to /h/ than the V1 curve. To find out whether there are significant quantitative differences in this direction, a Matlab-based procedure was designed and applied during this experiment (for the description of the procedure, see Section 3.4.7.1). Below, results from subjects S2 and S3 are presented. In the event, it was only possible to do quantification for these two subjects, because during the recording of subject S1, there occurred a slight shift forwards of the helmet with the transducer. Average distances between V1 and /h/ tongue contours and between /h/ and V2 tongue contours were calculated for S2 and S3. The distances are presented in Table 5-7. It is seen from the table that the distance from /h/ to the second vowel is not always smaller than the distance from the first vowel to /h/. In S2, there is a smaller C-V2 than V1-C distance in two vowel environments: /i/ and /a/. In S3, the smaller C-V2 distance is in the two high vowel environments.

Experiment 2: British English /h/

		ihi	uhu	aha
S2	V1-C	0.97	1.02	0.82
	C-V2	0.96	1.09	0.71
S3	V1-C	1.35	2.07	0.91
	C-V2	1.10	1.54	0.96

Table 5-7. Average distances in millimetres between V1 and /h/ and between /h/ and V2, for two subjects.

A Univariate ANOVA was conducted, separately for each subject, to explore the effect of the syllable affiliation of the vowel (V1 or V2) on the consonant-vowel distances, and the effect of the VCV type (/ihi/, /uhu/ or /aha/) on the consonant-vowel distances. The results for S2 show that there was no significant effect of the syllable affiliation of the vowel on the consonant-vowel distance. The results for S3 do show a significant effect: $F = 93.71$, $df = 1$, $p \leq 0.001$. Pairwise comparison for the S3 results was made, using the Bonferroni adjustment, and its results show a significant difference between the V1-C distance and the C-V2 distance, at the 0.05 level. A significant interaction was observed in both subjects, between the two independent variables, the VCV type (/ihi/, /uhu/ or /aha/) and the syllable affiliation of the vowel (V1 or V2): in S2, $F = 7.93$, $df = 2$, $p \leq 0.001$; in S3, $F = 46.91$, $df = 2$, $p \leq 0.001$. This interaction may account for the variation in the results for different vowel environments.

In both subjects, absolute distances are the greatest for the vowel /u/, and the smallest for the /a/ environment. The ANOVA results demonstrate a significant effect of the VCV type on the distances (S2: $F = 96.73$, $df = 2$, $p \leq 0.001$; S3: $F = 425.4$, $df = 2$, $p \leq 0.001$). The Games-Howell Post Hoc test shows significant differences between all pairs of VCVs (/ihi/ versus /uhu/, /ihi/ versus /aha/ and /uhu/ versus /aha/), for both subjects, at the 0.05 level. Some implications of these differences between vowels for our Hypothesis 2 will be discussed in Section 5.4.8.

The table also shows that the distances between /h/ and the vowels are always greater in S3 than in S2. A Univariate ANOVA was run, for the statistical comparison of the data from the two subjects. The results show that there was a significant difference between the subjects ($F = 432.38$, $df = 1$, $p \leq 0.001$). Pairwise comparison of the two subjects' results was made, using the Bonferroni adjustment, and its results show that the distances were significantly greater in S3 than in S2, at the 0.05 level. Individual differences in tongue position when producing target VCVs will be discussed in Section 5.4.7.

5.4. Discussion

5.4.1. V-on-C coarticulation in /h/

The experiment demonstrated that the tongue shape for /h/ varies according to the tongue shapes for the surrounding vowels. This result supports Hypothesis 1. This finding is similar to the results of Experiment 1, where bilabial consonants were shown to coarticulate greatly with the adjacent vowels. Interpreting the results in terms of the DAC model, we can say that the degree of resistance of /h/ to vocalic coarticulation in British English VhV sequences is lower than maximal.

In this experiment, a quantitative procedure was designed, that allowed for comparing sets of tongue contours for 15 repetitions of /h/ in different vowel environments. The procedure was applied in this experiment. It was shown that British English /h/ has significantly different tongue shapes when it occurs in the three different vowel contexts, /i/, /u/ and /a/. We can conclude that this procedure allows us to obtain quantitative data supporting Hypothesis 1, and to claim that there is a significant V-on-/h/ coarticulation.

5.4.2. Tongue movements in /ihi/ and /uhu/ sequences

Our results on /ihi/ and /uhu/ sequences did not support the hypothesis that there would occur a continuous tongue movement throughout the VhV sequence (Hypothesis 2). Qualitative analysis of whole tongue contours in V1, C and V2 showed that in /ihi/ and /uhu/ sequences, the tongue root was further back in C than in the two vowels, and the dorsum tended to be lower in C than in the vowels. These results were confirmed by the quantitative analysis of tongue displacement along a vertical measure bar: tongue lowering was observed between the two vowels in /ihi/ and /uhu/ sequences, in all the subjects, individually and pooled.

Experiment 2: British English /h/

In terms of the DAC model, troughs in English VhV sequences with high vowels can be considered a demonstration of CR by the intervocalic consonant. By suggesting that the consonant /h/ demonstrates some resistance to lingual coarticulation, it is implied that /h/ has its own identity in terms of lingual position. The results of this experiment, both from the whole contour analysis and from the measure bar calculations, suggest that the tongue dorsum in /h/ is lower than in the surrounding high vowels. The whole contour results suggest that the tongue root is retracted in /h/, in relation to the ajacent vowels /i/ and /u/. Then, within the DAC model, the interpretation of this lingual coarticulatory pattern is that the English consonant /h/ in VCV sequences has a supralaryngeal specification, with the tongue root being further back and the dorsum being lower than in the surrounding high vowels.

These data are consistent with some existing unpublished data from an experiment conducted by Maureen Stone with a group of colleagues (M. Stone, personal communication). It was a preliminary comparative ultrasound study of speech production, with one laryngectomy patient fitted with a Tracheo-Esophageal Puncture after surgery, and one control speaker. The subjects produced $C_1V_1C_1V_1C_1V_1$ syllables with five vowels (/i/, /a/, /u/, /æ/, /e/) and four consonants (/h/, /p/, /t/, /k/). One of the differences found between the patient and the control was that the tongue root during /h/ in /hæhæhæ/ sequences was more posterior to /æ/ for the control, and more anterior to /æ/ for the patient. The patient, as the researchers' interpretation suggests, was producing /h/ with a more anterior tongue root position because of a need to hold the cricopharyngeus muscle in a fairly open posture to produce less noise (in all the vowels the patient's tongue root was retracted, which suggested that the surgical modifications limited motion in the tongue root). The control speaker's retracting the tongue root for /h/ made the researchers suggest that in speakers with no disorders /h/ is specified for a rather posterior tongue root position.

Within the DAC model, it seems reasonable to suggest that the tongue root for producing British English /h/ is specified for a retracted position, in relation to the tongue root position for producing the vowels /i/ and /u/. The question arising is why the tongue root would be specified for a retracted position for producing /h/. A possible reason may be related to the larynx behaviour between the two high vowels. Contraction of the suprahyoid muscles involved in tongue raising for the high vowel production causes elevation of the larynx. Between the two vowels, these muscles would relax, this would result in larynx lowering, and would cause the tongue root to move backwards, and the dorsum to lower. It would be interesting to find an experimental approach that would give us more direct data on larynx behaviour, which would help to confirm or to reject the explanation proposed here.

This explanation is external to the DAC model, because it involves the larynx, and the DAC model is centered on tongue behaviour. The model can not offer an explanation in this case, it can only provide a way of registering the observed patterns in terms of CR. Specifying tongue root for a retracted position within the DAC framework arbitrarily reflects the physiological mechanism described above.

The results on lingual coarticulation in English /ihi/ and /uhu/ sequences are different from the results of the first experiment, on Russian VCVs with high vowels and bilabial consonants. There, no discontinuity in tongue movement between the two vowels was found (see Section 4.4.2 for discussion). This difference between the two experiments' results is interpreted in the DAC model terms by claiming that the tongue position is specified in different ways for Russian bilabial consonants and English /h/, therefore the coarticulatory patterns are different.

5.4.3. Tongue movements in /aha/ sequences

The results for /aha/ sequences supported the hypothesis that there would occur a continuous tongue movement throughout the VhV sequence (Hypothesis 2). Qualitative observations of whole tongue contour sequence patterns did not show any strong tendencies for discontinuity in coarticulation. Measure bar results did not show significant evidence towards discontinuity in coarticulation, for the three subjects pooled. Individually, the subjects tended to differ in their ways of producing the /aha/ sequence. However, these differences can be explained by transducer orientation under the chin (see more on this in Section 5.4.8).

Experiment 2: British English /h/

Within the DAC framework, the same interpretation could be offered as the one used for the high vowel contexts. It was suggested in Section 5.4.2 that English /h/ is specified for a retracted tongue root position, in relation to the vowels /i/ and /u/. In /aha/ sequences, the tongue root was further back in our data than in /ihi/ and /uhu/ sequences, and the tongue root position was comparable in the C curve and the two vowel curves. It is possible that no significant discontinity in coarticulation occurred in /aha/ sequences because the specification of the tongue root in /h/ was closer to the specification of the tongue root in the surrounding vowels than in the case of /ihi/ and /uhu/ sequences. For producing the vowel /a/, the tongue is already low and has a rather flat posture (i.e., the tongue root is retracted), so there should not be much displacement for the consonant. This interpretation is consistent with the explanation for the discontinuity in lingual coarticulation found in /ihi/ and /uhu/ sequences (Section 5.4.2). For producing the low vowel /a/, the larynx is not raised, and hence syprahyoid muscles do not relax between the two low vowels.

These results are different from the results of the first experiment. There, a significant number of antitroughs – tongue dorsum raising patterns between the two vowels – was found. This difference between the two experiments' results can be interpreted in the DAC model terms by suggesting that the tongue position is specified in different ways for Russian bilabial consonants and English /h/, and so the coarticulatory patterns are different.

5.4.4. Alternative interpretations of the observed tongue movements in /ihi/, /uhu/ and /aha/ sequences

In Sections 5.4.2 and 5.4.3, arguments were presented for the explanation of the observed patterns by CR, i.e., tongue position specification for the consonant. In this section, main alternative explanations to the observed patterns and the reasons for rejecting these explanations are discussed.

The mechanism responsible for discontinuity in lingual position in VCV sequences proposed by Lindblom et al. (2002) was described as "a momentary deactivation of the tongue movement after V1" (Lindblom et al. 2002, p. 245). In our data, tongue lowering occurring in /i/ and /u/ contexts could be interpreted by suggesting that the tongue deactivates its position required for the vowel and moves towards a more neutral, schwa-like position for the consonant. A continuous tongue movement in /a/ contexts found in our data cannot be regarded as tongue deactivation between two /a/ vowels. So the tongue deactivation mechanism cannot be taken as fitting the results of lingual coarticulation in /aha/ sequences in the present study.

Another interpretation of the observed patterns could involve an aerodynamic effect. Some explanations of lingual coarticulatory patterns in VCVs with bilabial consonants by aerodynamics have been offered in the literature (see Sections 2.2.4, 2.2.6, 2.2.7, 2.2.8). Of course, in the case of the glottal fricative, unlike bilabial stops, the aerodynamic effect could be almost discounted, because there is no occlusion, and hence no pressure rise during the consonant production. However, there might be a slight effect due to devoicing between the vowels, which increases airflow. For example, in /ihi/ and /uhu/ sequences, increased airflow, due to devoicing during the /h/ production, would arguably result in tongue lowering. But the continuous tongue movement in /aha/ sequences could not be explained by this aerodynamic effect; so aerodynamics do not fully explain our data.

Another possible explanation of the tongue behaviour in VhV sequences could be the influence of the suprasegmental organisation of the VCV, i.e., the requirement to produce two consecutive syllables. This explanation is reminiscent of the claim made in Gay (1977) that the CV sequence is “organized and produced as an integral articulatory effect” (Gay 1977, p. 192), based on the findings of Kozhevnikov and Chistovich (1965) and on his own findings, and of the suggestion by Perkell (1986) about the syllable organisation of motor commands (see also Section 2.2.3). This interpretation would suggest that at some point during the VCV there occurs a gap, or a trough, in the muscular effort, required to pass from producing one syllable to producing another syllable. This gap in the muscular effort arguably results in discontinuity in the lingual movement throughout the VCV sequence. Then the intervocalic tongue lowering in /ihi/ and /uhu/ sequences could be involved in the mechanism of making the two syllables of the V#hV sequence audibly distinct from each other. However, the absence of discontinuity in tongue movement throughout /aha/ sequences demonstrated in this experiment does not allow us to take the “syllabic” explanation as accounting for all our data.

One more way of explaining some points of our results concerns vowel diphthongisation. In Section 2.2.5, the study by Perkell (1986) was described, which claims that discontinuity in coarticulation in VCV sequences may arise from producing the vowels of a VCV sequence as diphthongs. Diphthongisation was also suggested by Bryan Gick (personal communication) as a possible explanation of the trough pattern found in VhVs with high vowels in this experiment. There was some diphthongisation in our data, as evidenced by auditory analysis. For example, S3, unlike the other subjects, produced both /u/ vowels of the /uhu/ sequence as diphthongs. Typical pronunciations of /uhu/ sequences in the three subjects could be transcribed as follows: [ʉ̟xʷʉ̟] in S1, [ʊːxʷʊː] in S2, and [ʏʉxʷʏʉ] in S3. In /ihi/ sequences, all the subjects produced both vowels as monophthongs. In Section 5.3.1 (Figure 5-6), we observed that in /uhu/ sequences, the consonant curve was further away from both vowel curves in S3, than in the other two subjects. In Section 5.3.6, we obtained comparable results for the measure bar distances: in S3, unlike the other two subjects, both tongue displacements (from V1 to C and from C to V2) in /uhu/ were significantly greater than in the other two vowel environments. This might be a result of a diphthong-like production of /u/ in this subject, compared with S3’s other vowels and with the other two subjects’ productions. But then we need an explanation for troughs found in other subjects, and also for troughs in /ihi/ sequences. These cannot be explained by diphthongisation, so we cannot take it as an explanatory mechanism for all our data.

5.4.5. Comparison of tongue movements in VCV sequences with non-lingual consonants, across language and across consonant type

In Section 4.4.5, a cross-linguistic comparison was made of tongue behaviour in VCVs with bilabial consonants. Here, British English /h/ is discussed in relation to the bilabial stops. Three sets of results are compared: Russian VCVs with bilabial stops (Experiment 1), British English VCVs with bilabial stops (Vazquez Alvarez et al. 2004), and British English VhV sequences (Experiment 2).

The distribution of the four tongue contour sequence patterns in the three sets of data is presented in Table 5-8. In the Russian data for bilabial stops, there were 90 tokens in each of the three vowel groups. In the English data for /h/, there were 45 tokens in each of the three vowel groups. In the English data for bilabial stops, there were 100 tokens in each vowel group. Percentage values are presented in the table, to facilitate the comparison.

	/i/			/u/			/a/		
	Russian bilabial stops	*English /h/*	*English bilabial stops*	*Russian bilabial stops*	*English /h/*	*English bilabial stops*	*Russian bilabial stops*	*English /h/*	*English bilabial stops*
Trough	31	91**	86**	9**	96**	76**	**	24	33
Anti-trough	3**	**	3**	20	**	1**	67**	36	34
Contin. up	66**	2**	5**	71**	2**	**	33	31	19
Contin. down	**	7*	6**	**	2**	23	**	9*	14*

Table 5-8. Rate of occurrence (percentage) of tongue contour sequence patterns, in three sets of data: Russian VCV sequences with bilabial consonants; British English VhV sequences; British English VCV sequences with bilabial consonants. The data from all subjects are pooled and presented according to vowel type. Two black asterisks next to a number mean that the rate of occurrence was significantly above chance, at $p < 0.001$. Two red asterisks next to a number mean that the rate of occurrence was significantly below chance, at $p < 0.001$. One red asterisk next to a number means that the rate of occurrence was significantly below chance, at $p < 0.01$. For details about statistical calculations, see Appendix IV.

It can be seen from the table that the distribution of tongue contour sequence patterns is very close in the two sets of English data – bilabial stops and /h/. Both of these data sets are different from the Russian data on bilabial stops. In Section 4.4.5, the difference in tongue behaviour between Russian and English bilabial stops was explained within the DAC model, by a different tongue position specification in these two languages. As it was shown in Experiment 2 (Sections 5.4.2 and 5.4.3), the lingual coarticulatory pattern observed in English VhV sequences could be explained by the tongue root position specification for producing the intervocalic /h/. It is interesting that the coarticulatory patterns in English bilabials and /h/ are so similar, and that both of them are different from the coarticulatory pattern in Russian bilabial consonants. This cross-linguistic difference suggests that CR specifications for non-lingual consonants are more similar within one language than across languages. If larynx lowering between two high vowels is the reason for discontinuity in lingual coarticulation, as suggested in Section 5.4.2, then the same explanation could be offered for the lingual behaviour in British English VCVs with high vowels and bilabial stops. Russian bilabial stops, unlike English non-lingual consonants, have a constraint on tongue position, and hence, on larynx position, for their production, so there is no relaxation of suprahyoid muscles between two high vowels in Russian.

5.4.6. Syllable boundary influence on VhV coarticulation

Similarly to Experiment 1, some qualitative observations were made about the V1-C and C-V2 distances. In order to show quantitatively whether the C-V2 distance was consistently smaller than the V1-C distance (Hypothesis 4), a Matlab-based procedure was developed. This procedure allowed for measuring the distance between the sets of curves for the consonant and the sets of curves for each vowel, and to compare these two distances for significant difference. The procedure was applied in this experiment, for the data from two subjects. The results showed that only in one subject the C-V2 distance was significantly smaller than the V1-C distance. So Hypothesis 4 is only partly supported by these results. A possible reason for the difference between the subjects may be that the three curves (V1, C and V2) were very close together, and there was not much opportunity for the V1-C distance to be significantly greater than the C-V2 distance. In S3, both distances were significantly greater than in S2 (see Section 5.3.7), and S3 was the subject who had a significant difference between the two distances.

Interpreting the results in terms of the DAC model, we can say that in our data the consonant is not always less resistant to the vowel in the same syllable than to the vowel from a different syllable. So the DAC model cannot treat our data as evidence that the DAC value of the intervocalic consonant in V#hV sequences is always influenced by the syllable affiliation of the consonant.

In Chapter 7, the findings discussed in this section and the claims formulated about these findings will be developed. Also, suggestions for quantifying the DAC values of the intervocalic /h/ will be presented, with examples of calculations, based on the ultrasound data from this work.

5.4.7. Inter-subject differences

In this experiment, some inter-subject differences can be explained by the vocal tract morphology or the transducer placement under the chin. For example, in Figures 5-5 – 5-7, we can notice differences in individual shapes of the subjects' tongues. This is natural and expected, and it was also shown in Experiment 1 (see Section 4.3.7).

Some differences between subjects observed in this experiment are explained by the position and angle of the transducer. For example, in Section 5.3.4, a difference was observed between subject S1 and the other two subjects, in that S1 produced a significant number of troughs in /aha/ sequences, while the other two subjects did not. Also, in Section 5.3.6, it was noted that patterns of tongue displacement in /aha/ sequences based on the measure bar results were different across subjects. The differences in tongue displacement were very small in all the subjects, because displacement itself was small in the /a/ environments. Together with the differing probe orientation under the chin, this contributed to the cross-subject difference in the results (more details about probe angle in relation to the interpretation of our results are presented in Section 5.4.8).

Some individual differences were described in Section 5.3.1. Qualitative analysis of whole tongue contours showed that in S3, the contours for V1, C and V2 were not as close together as in the other two subjects. Supporting evidence for this claim comes from Section 5.3.7, where V1-C and C-V2 distances were presented for the subjects S2 and S3. It was shown that the distances between /h/ and the vowels were always significantly greater in S3 than in S2. This could be explained by suggesting that in S3, the /h/ in VhV sequences is coarticulated less than in S2. In terms of the DAC framework, we can interpret this by claiming that the DAC values of /h/ in S3 are greater than in S2.

Another difference between S2 and S3 was reported in Section 5.3.7 and discussed in Section 5.4.6. In S3, the C-V2 distance was significantly smaller than the V1-C distance, but not in S2. This was interpreted by saying that in V#hV sequences, the resistance of the intervocalic /h/ to V2 is not necessarily smaller than its resistance to V1. Here, we can suggest that the resistance of /h/ to the syllable boundary influence can vary across subjects.

One more inter-subject difference again demonstrates an individual strategy of S3 in VhV production. In this subject, unlike the other two subjects, a significant difference in both tongue displacements (V1-C and C-V2) between /u/ and the other two vowels was reported in Section 5.3.6. This may be explained by the diphthong-like production of /u/ in this subject, as compared with S3's other vowels and with the production of the other two subjects. Typical pronunciations of /uhu/ sequences in S3, as noted in Section 5.4.4, could be transcribed as [ʏʉxʷʏʉ].

5.4.8. Methodological issues: calculations based on a measure bar

In Section 4.4.9, it was discussed how the transducer orientation under the chin could have influenced the results of the analysis based on the highest point in the C curve. In this experiment, in addition to the method used in Experiment 1, quantitative analysis of tongue displacement was used, based on a vertical measure bar. So the question of transducer orientation is important here, too.

It was shown in Section 4.4.9 that in Experiment 1, the angle between the probe and the chin did not affect most results of the calculations based on the highest point in the C curve. In this experiment, there were more differences in the transducer orientation between the subjects than in Experiment 1. The ultrasound scans for subjects S2 and S3 (Appendix V-2) demonstrate that the transducer is at a different angle under the chin in these two subjects. We can be certain about the orientation of the transducer by looking at the position of the chin and hyoid bone shadows in relation to the horizontal plane (see Figure 4-2 for an example of the symmetrical location of the two shadows, representing a right angle between the transducer and the line of the jaw). In S2, the transducer is rotated anti-clockwise, while in S3, the rotation is clockwise. It can also be seen in Appendix V-2 that in S1, there is some transducer rotation clockwise, but less than in S3. This difference in transducer orientation can also be seen in Matlab plots (for example, compare Figures 5-12b and 5-12c, with the graphs for the /aha/ sequence in the subjects S2 and S3, respectively).

These differences in the transducer position in relation to the chin have implications when it comes to determining the highest point of the tongue. Therefore, it is important to compare the results of the measure bar calculations with the results of the qualitative analysis of whole contours.

In Section 5.3.1, we made a qualitative observation that the three tongue contours in /aha/, for V1, C and V2, were very close together. This observation was confirmed by the measure bar results (Sections 5.3.5 and 5.3.6): tongue displacements from V1 to C and from C to V2 were significantly smaller in /aha/ sequences than in /ihi/ and /uhu/ sequences, across subjects. The calculation of average distances between the curves produced similar results: the V1-C and C-V2 distances were significantly smaller in /aha/ than in /ihi/ and /uhu/ sequences (Section 5.3.7). So, considering this small extent of tongue displacement in /aha/ sequences, it is clear that even a small rotation of the ultrasound transducer could have affected the results based on the measure bar. That is why we observed differences in individual results based on measure bar calculations. However, the qualitative comparison of whole tongue curves did not show any strong discontinuity in tongue movement from V1 to V2, for each subject. This allows us to conclude that in /aha/ sequences, there was a continuous tongue movement from V1 to V2 throughout the consonant /h/.

As stated in the previous paragraph, in /ihi/ and /uhu/ sequences, the tongue displacements from V1 to C and from C to V2 were greater than in /aha/ sequences. The particular displacement patterns of individual parts of the tongue in /ihi/ and /uhu/ sequences (namely, retracting the tongue root between the vowels, lowering the dorsum, and slightly retracting the blade) were described in Section 5.4.2. Qualitative analysis of whole tongue contours in /ihi/ and /uhu/ sequences made it possible to estimate whether the results of the measure bar calculations corresponded to the patterns of displacement of parts of the tongue. This analysis showed that the measure bar results did reflect the actual displacement pattern, even though the tongue curves had different orientation in different subjects. Should the transducer orientation have been the same across subjects, the results of the measure bar procedure in /ihi/ and /uhu/ sequences would have been comparable to the results obtained in this experiment, namely, a significant number of troughs.

5.4.9. Relative importance of parts of the tongue in describing tongue movements

In Section 4.4.10, some discussion was offered of possible specification of parts of the tongue, in relation to the whole tongue. Here, this discussion is continued, based on the data obtained in Experiment 2.

In this experiment, troughs have been observed in British English VhV sequences with high vowels. A trough was defined as a displacement of the tongue downwards, along a vertical measure bar imposed on the highest point in the C curve. This displacement was called tongue lowering. Qualitative analysis of whole contours (Figures 5-5 – 5-7) shows that the tongue movement cannot be entirely described by the word “lowering”. The data from the whole tongue contour in /ihi/ and /uhu/ sequences show a common pattern of tongue behaviour during the consonantal portion of the VhV. The tongue root in all the subjects is further back for the consonant than for both vowels, the dorsum is generally lower for the consonant than for the vowels, and the blade is either below or between both vowel curves.

This pattern of tongue movement clearly shows that the root, the dorsum and the blade behave differently. This description of the whole tongue contour changing throughout the VCV does not clarify whether the whole tongue body is equally active in this movement between the two vowels, or whether some part(s) of it just passively follow(s) others. For example, the tongue root may be actively moving backwards, causing concomitant dorsum lowering, and also some adjustments of the blade position. This is what has been suggested in this experiment as an interpretation of the observed pattern within the DAC model (see Section 5.4.2). However, it is rather difficult to think of a way to demonstrate that this is the mechanism governing the tongue movement. One of the indications that the root is the specified tongue part may be the following. In S2, average tongue curves for /aha/ (Figure 5-7b) show that the very back part of the root is further back in the C curve than in the vowel curves. In S2, this pattern was rather consistent across repetitions. This may be an indication that S2 was retracting the tongue root for producing the intervocalic /h/, even though both low vowels in /aha/ sequences were produced by this subject as back vowels ([ɑ]). In S1, the average tongue curve pattern (Figure 5-7a) also shows a very small retraction of the back part of the tongue root between the vowels; in this subject, however, this pattern was not as consistent across individual repetitions as in S2. In subject S3, the average tongue contour pattern in /aha/ (Figure 5-7c) is different from that of both other subjects, because S3 produced /aha/ sequences with the stress on the second syllable. The first, unstressed, vowel was consistently realised by S3 as a schwa, and the second vowel was [ɑ]. The other two subjects realised both syllables with approximately equal stress and both vowels had a similar quality.

At this stage, it seems appropriate to suggest a way of measuring discontinuity in lingual coarticulation based on displacement of parts of the tongue, and not just a point-based measurement. Given our suggestion that in /h/, tongue root is specified for lingual position, we could measure the displacement of the root separately from the rest of the tongue. As suggested earlier (Section 5.4.2), in VCV sequences with high vowels, according to our data, the tongue root is retracted, it pulls the rest of the tongue, and this results in the lowering of the dorsum, in relation to its position during the vowels. The point-based measurements in this work were based on tongue displacement along a vertical line imposed on the apex of the consonant curve, so the trough pattern was observed by registering tongue dorsum lowering. Calculations of tongue root displacement, as our results suggest, would be a more direct way of measuring discontinuity in lingual coarticulation in VhV sequences with high vowels.

An idea for measuring displacement of tongue root in /ihi/ and /uhu/ sequences is as follows. The V1-C distance and the C-V2 distance in the region of the tongue root would form a basis for quantifying discontinuity in tongue movement. Averaging these two distances could give us the size of a trough in tongue movement, based on the tongue root displacement.

An interesting question is whether this tongue displacement would be detectable with other articulatory techniques, for example, EPG. No EPG data were collected in this experiment, so it is impossible to give a confident answer. Considering that existing studies have detected troughs in tongue movement in /ipi/ sequences with EPG (e.g., McAllister & Engstrand 1991), we could speculate that a similar discontinuity in the amount of tongue contact with the palate could be traced using EPG. However, ultrasound appears to be a more suitable technique for measuring this discontinuity. As we have seen in this experiment, troughs occur not only in the middle part of the tongue, but also in the tongue root. EPG could capture the pattern of contact of the middle part of the tongue with the palate, but not the root movement.

Continuing the discussion from Section 4.4.10, we will make some further comments about displacement of parts of the tongue, in relation to the whole tongue contour. The data presented in this experiment suggest that we can talk about the tongue being back or forward when the back and front thirds of the tongue are displaced along the x axis to the left or to the right, respectively. The tongue is described as being higher or lower when the middle third of the tongue is displaced along the y axis, upwards or downwards, respectively. This description of displacement of the back and front parts of the tongue mostly along the x axis, and the middle part of the tongue mostly along the y axis is dependent on the structure of the tongue, and its location in the mouth. Unlike the jaw or velum, the whole tongue body displacement cannot be properly measured along a “high”-“low” scale, or “front”-“back”. The tongue body is pulled by numerous muscles, and it changes its position along three dimensions at the same time. Thus, measuring the displacement of the whole tongue along any of these two scales is inappropriate. When the tongue is in the resting position, it has a shape close to a semicircle (see Appendix V). So it is natural to describe movements of the tongue parts in Cartesian space in the way outlined here.

5.5. Summary

In this experiment, coarticulatory patterns in British English VhV sequences were studied. The aims were to test some claims from the literature about coarticulatory characteristics of /h/, and to add more information to the results of the first experiment, where the behaviour of bilabial consonants was analysed. The notion of CR was used in formulating the hypotheses and interpreting the results.

A significant number of troughs in /ihi/ and /uhu/ sequences and a continuous tongue movement in /aha/ sequences were reported. These findings were interpreted within the DAC model by claiming that the tongue has a lingual position specification for the production of the British English consonant /h/. Tongue root was suggested to be specified for a retracted position in this consonant.

The hypothesis that the intervocalic /h/ in V#hV sequences would be less resistant to V2 than to V1 was not fully supported by the experimental data. Productions of two subjects were quantitatively analysed. In one subject, this hypothesis was supported, and in the other subject, it was not. The results were interpreted as suggesting that in V#hV sequences, the resistance of /h/ to syllable boundary influence can vary across subjects.

Some more inter-subject differences were reported in the experiment, that could not be explained by transducer position or individual vocal tract configurations. These differences were interpreted within the DAC model, by saying that the CR of speech sounds varies across speakers.

Experiment 2: British English /h/

In this experiment, two quantitative methods of tongue curve comparison were designed and applied, that made it possible to solve the problems formulated in the process of Experiment 1 (see Sections 4.4.1 and 4.4.7). One procedure allows for defining whether two sets of curves are significantly different from each other. Another procedure allows for claiming, based on whole tongue contours, whether one tongue displacement is significantly different from another tongue displacement. These procedures will be used for data analysis in Experiment 3.

Quantitative data from the whole tongue curves obtained in this experiment will be used in Chapter 7, where the degree of CR in /h/ will be quantified and compared to that of lingual consonants and vowels.

6. EXPERIMENT THREE: COARTICULATION RESISTANCE IN BRITISH ENGLISH LINGUAL STOPS AND VOWELS

6.1. Introduction

In Experimens 1 and 2, resistance to lingual coarticulation was studied in non-lingual consonants. Bilabial stops and the glottal fricative were suggested to have some resistance to the surrounding vowels in a VCV sequence. However, a very significant amount of coarticulation was found in those consonants. In this experiment, the question of how resistant speech sounds are to the influence of neighbouring sounds was pursued. More data and more advanced methods are involved. The purpose of this experiment was to test some claims formulated within the DAC model, and to add new information towards a more accurate formulation of articulatory constraints.

The concept of Coarticulation Resistance (CR) has to do with measuring and predicting the scope that speech sounds have to adapt to other sounds, without losing their identity (i.e., without losing t he possibility of being correctly perceived by listeners). So theoretically, there can be two extreme cases: one where the sound totally adapts to the neighbouring sound(s), and thus has a zero DAC value, and one where the shape of the articulators during production of a sound in two different contexts does not change at all in respect to the neighbouring sounds, and thus this sound has an absolute DAC value. This experiment was centered on measuring coarticulation resistance in lingual consonants and in vowels. Based on the literature (e.g., Keating et al. 1994; Recasens et al. 1997; Recasens 1999), it was expected that these sounds would exhibit higher DAC values than non-lingual consonants analysed in Experiments 1 and 2. The DAC model is heavily based on data from VCV sequences (for more details, see Section 2.1.2), so choosing VCV sequences to analyse with ultrasound provides more grounds for comparing my results to the earlier published results within the DAC model.

There do not exist ultrasound studies of coarticulatory influence of vowels on lingual consonants within the CR approach. The articulatory data presented in Recasens et al. (1997) were obtained using EPG. Before embarking on the ultrasound analysis in this study, it was important to make sure that similar types of results could be obtained to those presented in Recasens et al. (1997), using a similar articulatory technique to the one they used.

In this experiment, a multi-channel technique was introduced, which allows for simultaneous ultrasound, EPG and acoustic data collection and analysis. In the QMUC ultrasound system, there exists the possibility of recording the EPG signal combined and synchronised with ultrasound and acoustics (see Section 3.3 for details). Ultrasound and EPG data were obtained for demonstrating V-on-C coarticulation. Collecting both types of data made it possible to support ultrasound results with the information on the exact places on the hard palate touched by the tongue, and to be able to more directly compare the results of this experiment to those presented in the studies within the DAC model.

To the best of my knowledge, there are very few studies using this combined methodology for speech research. In Stone et al. (1992), cross-sectional tongue shapes and linguopalatal contact were examined in American English CVC utterances with the consonants /s/, /ʃ/, /l/, and the vowels /i/, /ɛ/, /o/ and /ɑ/. In that study, “the tongue-palate contact patterns were used to provide a better understanding of how different tongue shapes are produced by identifying the parts of the palate against which the tongue might be braced” (Stone et al. 1992, p. 254). No midsagittal scans of the tongue were collected in Stone et al. (1992), only coronal scans were analysed. Stone and Lundberg (1996) used ultrasound in order to reconstruct three-dimensional tongue surfaces from the production of various sustained American English consonants and vowels. EPG data were also collected by these researchers, in order “to compare tongue surface shapes with tongue-palate contact patterns” (Stone & Lundberg 1996, p. 3728). No acoustic data were used in that study. Cohen et al. (1998) used EPG and ultrasound as supporting data for their visual speech synthesis. Scobbie et al. (2004) was a methodological description of the possible use of ultrasound and EPG together.

In Figure 6-1, linguopalatal configurations are presented for several lingual consonants from /iCi/ and /aCa/ sequences, investigated in Recasens et al. (1997). These researchers showed, based on the data from five repetitions of VCV sequences, that several Catalan consonants differed systematically, depending on the vowel context, /i/ versus /a/. One of the goals of the EPG experiment in this work was to confirm, using EPG data, that there occurs a significant effect of the vowels on the intervocalic consonant. Statistical calculations were performed on the linguopalatal contact data for the consonant /t/ from /iti/ and /ata/ sequences, in order to show whether the consonant lingual contact patterns are significantly different in two different vowel environments, i.e., whether the V-on-C effect is observed in the EPG data. The consonant /t/ was chosen for the EPG experiment because it was not included in Recasens et al. (1997). This was therefore an opportunity to extend the picture Recasens and his colleagues presented in their work.

An ultrasound analysis of the V-on-C effect was conducted. Tongue curves for /t/ from /iti/ and /ata/ sequences were compared. Statistical analysis was carried out, to show whether the lingual shapes for /t/ in two different vowel environments differ significantly from each other, i.e., whether the V-on-C effect is observed in the ultrasound data.

C-on-V coarticulation was measured with ultrasound, in order to obtain numerical evidence of the presence/absence of an effect. There was the same method of measuring C-on-V coarticulation as the one used to measure V-on-C coarticulation. Statistical analysis determined whether the lingual shapes for /a/ in two different consonant environments differed significantly from each other. For measuring the C-on-V effect, tongue curves for the vowel /a/ from /ata/ and /aka/ sequences were compared.

By using the same method for measuring V-on-C and C-on-V coarticulation, it was possible to directly compare CR of consonants and vowels, based on ultrasound results. The sizes of V-on-C and C-on-V effects were compared. Based on the existing literature (e.g., Öhman 1966; Keating et al. 1994), we could expect that vowels will be less influenced by consonants than consonants by vowels.

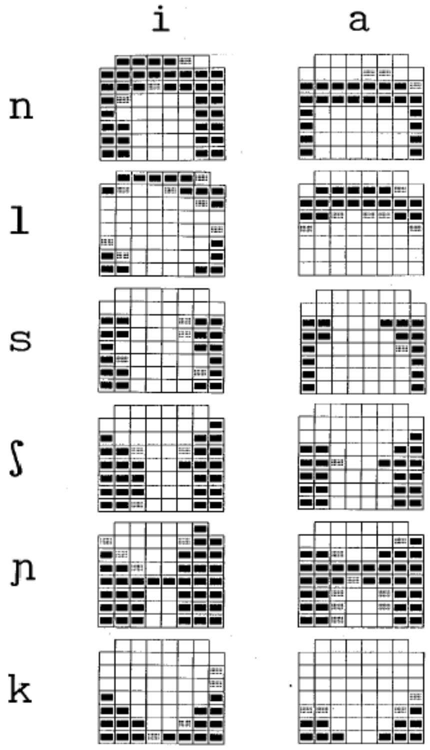


Figure 6-1. Linguopalatal configurations for several Catalan lingual consonants from symmetrical VCV sequences with the vowels /i/ and /a/ (after Recasens et al. 1997). The measurements were taken at the consonantal midpoint. The data for each consonant are based on five repetitions. Black: 80-100% electrode activation; grey: 40-80% electrode activation; white: less than 40% electrode activation.

The differences between V1-C and C-V2 distances were examined, in order to investigate whether the syllable boundary has an effect on the consonant's resistance to the influence of the surrounding vowels. The syllable boundary influence was a predicted pattern in this experiment.

6.1.1. Hypotheses

The experiment was designed to test the following hypotheses:

1. **EPG patterns for /t/ in the context of two vowels, /i/ and /a/, will provide evidence in support of a V-on-/t/ effect: specifically, there will be a significant difference between the contact pattern for /t/ in the two vowel environments.** If the hypothesis is supported, it will be concluded that /t/ is influenced by the neighbouring vowels. This hypothesis will be refuted if there is no significant difference between the contact pattern for /t/ in the two vowel environments. On the basis of this result, it will be concluded that the DAC value of /t/ in the context of these vowels is the highest possible.

2. **There will be a V-on-/t/ effect in the sequences /ata/ and /iti/: specifically, there will be a significant difference between the ultrasound tongue contours for /t/ in the two vowel environments.** If the hypothesis is supported, it will be concluded that /t/ is influenced by the neighbouring vowels, and that its DAC value in VCV sequences is lower than maximal. This hypothesis will be refuted if there is no significant difference between the tongue contours for /t/ in the two vowel environments. If this result occurs, it will be concluded that the DAC value of /t/ in the context of these vowels is the highest possible.

3. **There will be a C-on-/a/ effect in the sequences /aka/ and /ata/: specifically, there will be a significant difference between the ultrasound tongue contours for /a/ in the two consonant contexts.** If the hypothesis is supported, it will be concluded that /a/ is influenced by the neighbouring consonants, and that its DAC value in VCV sequences is lower than maximal. This hypothesis will be refuted if there is no significant difference between the tongue contours for /a/ in the two consonant contexts. On the basis of this result, it will be concluded that the DAC value of /a/ in the context of these consonants is the highest possible.

4. **The V-on-/t/ effect will be significantly greater than the C-on-/a/ effect.** If the hypothesis is supported, it will be concluded that the /t/ is influenced by the neighbouring vowels more than the /a/ is influenced by the neighbouring consonants, and that consequently, in these particular contexts, the DAC value of /a/ is greater than

that of /t/. This hypothesis will be refuted if there is no significant difference between the V-on-/t/ effect (measured on the ultrasound data) and the C-on-/a/ effect, or if the C-on-/a/ effect is greater than the V-on-/t/ effect. These would mean, respectively, that there is no significant difference between DAC values of the /t/'s and the /a/'s in these contexts, or that the DAC value of /t/ is greater than that of /a/.

5. **In a V#CV sequence (where "#" signifies word boundary), there will be a significant difference between the V1-C distance and the C-V2 distance: specifically, the distance between the V1 curve and the C curve will be significantly greater than the distance between the C curve and the V2 curve.** If this hypothesis is supported, it will be concluded that there is a word boundary influence on the consonant resistance to the surrounding vowels, and that the consonant is less resistant to V2 than to V1. This hypothesis will be refuted if there is no significant difference between the V1-C distance and the C-V2 distance, or if the V1-C distance is significantly smaller than the C-V2 distance. These results would mean, respectively, that the consonant's resistance to the surrounding vowels is not influenced by the word boundary in V#CV sequences, or that the consonant is more resistant to V2 than to V1.

6.2. Method

6.2.1. Experimental items

The stimuli were the sequences /a#ta/, /i#ti/, and /a#ka/ ("#" signifies word boundary here), occurring in the following meaningful sentences: "At 4 pm Ma tasked Janet to paint the roof"; "Little Leigh teased Janet"; "After that Ma cast an angry look at Leigh". The three sequences were embedded in the following words: "M**a ta**sked", "L**eigh tea**sed", "M**a ca**st". The text in bold underlined script shows the target sequences: /a#ta/, /i#ti/, and /a#ka/. It was predicted that the syntactic structure of the sentences would prompt the subjects to produce the two target vowels with equal stress, and, as expected, all the subjects produced two more or less equally stressed vowels in the VCV stimuli.

6.2.2. Subjects

There were three subjects, all native speakers of English. They all had a Southern British accent, and similar acoustic characteristics of the vowels used in the stimuli. Two speakers wore their artificial EPG palates. One speaker did not wear the artificial palate for technical reasons. However, this was not a problem, because the aim of the EPG experiment was largely methodological, and data from two speakers were sufficient for deciding whether the EPG results were informative and reliable.

6.2.3. Instrumentation and recording procedure

The details of the recording procedure in respect of the ultrasound are as described in Chapter 3 (Section 3.4.1). In this experiment, EPG data were also recorded from two speakers, together with the ultrasound and acoustic signals.

There were 15 tokens in each stimulus type. The total number of VCV sequences recorded and analysed in this experiment was 135 tokens.

The order of presentation was the following. One repetition of each of the three sentences was collected as a block, before moving on to the second block, and so on. The sentences in each block were presented in the same order. The order was: “At 4 pm Ma tasked Janet to paint the roof”; “Little Leigh teased Janet”; “After that Ma cast an angry look at Leigh”.

The participants were given a printout of the sentences, for some pre-recording practice. The subjects were asked to produce the sentences at a comfortable speaking rate.

The two speakers who wore the EPG palates were experienced EPG users. The palates were checked before the experiment, to ensure that there were no loose electrodes. The subjects were instructed to insert their palates before putting the ultrasound helmet on. Setting up ultrasound and EPG at once required more care than preparing for a customary ultrasound recording, in order to make the subjects feel comfortable, so it took slightly more time than preparations for an ultrasound recording alone.

6.2.4. Ultrasound software analysis, annotations and splines

Annotating the waveform and creating splines within the ultrasound analysis software (Articulate Assistant), common to all three experiments in this work, was described in Section 3.4.4. An illustration of the annotations and spline drawing in this Experiment is given in Figure 6-2. In this experiment, the V1 spline was placed at the mid-point of V1 (note that in Figure 6-2, the vowel is preceded by the nasal consonant /m/), the C spline was placed at the mid-point of the stop consonant closure, and the V2 spline was created at the V2 annotation point described in Section 3.4.4.

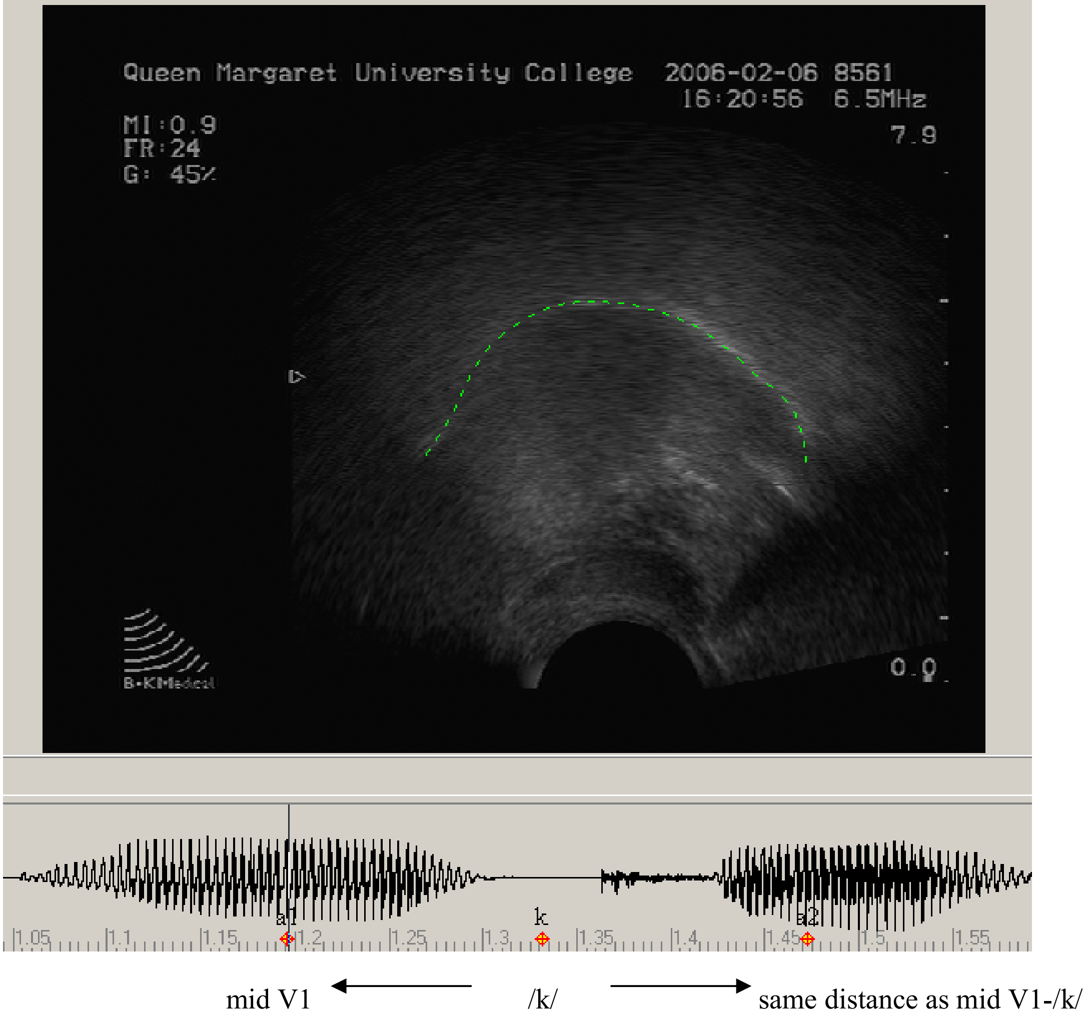


Figure 6-2. Illustration of the three annotation points and a spline drawn at the tongue contour, corresponding to the /a1/ annotation point (/a1/ means the mid-point of the first vowel).

After annotating the waveform and drawing the splines, the xy data for each curve were exported from Articulate Assistant into a text file, and then imported into Matlab for analysis. Calculations in Matlab were done for each speaker separately. Matlab calculations specific to this experiment are described in detail below (for the common Matlab procedures described earlier, in Chapter 3, references to relevant sections will be provided where necessary).

6.2.5. Qualitative observation of whole tongue contours

Some qualitative analysis of the curves was made, together with the quantitative analysis. Two different types of plots were used for portraying whole tongue contours. One of them consisted in plotting on the same graph average tongue curves for V1, C and V2 belonging to the same VCV sequence (see Section 3.4.5 for the details of the averaging procedure and for an example of this type of graph). Another way of displaying tongue curves involved plotting two sets of 15 curves, representing the same sound in two different contexts (see Section 3.4.5).

6.2.6. V-on-C coarticulation, EPG data analysis

V-on-C coarticulation was measured by comparing the Q_p index during the consonant closure for the consonant /t/, in the context of two vowels, /i/ and /a/. The Q_p index is a measure used in Recasens et al. (1997) for quantifying coarticulatory effects using EPG. Q_p represents the percentage of contact activation over the palatal zone, i.e., the ratio of the number of activated palatal electrodes to the total number of palatal electrodes. When the index is applied to measuring vocalic influence on intervocalic consonants, it quantifies the amount of vowel-related activation of the tongue dorsum that is present during the consonant production.

The palatal zone is represented by the three back rows of the artificial palate. This functional division of the artificial palate is used in Recasens et al. (1997) for calculating the Q_p index for alveolar consonants, in order to represent the degree of vowel-related tongue dorsum raising during consonant production. There exist alternative views on the functional division of the artificial palate (see, e.g., Wood 1997). In this work, the criteria defined by Recasens and his colleagues are used, in order to be able to compare our results to the results presented in Recasens et al. (1997). For the sake of preserving this comparability, other existing EPG data reduction methods were not used in this experiment (e.g., coarticulation index devised by Farnetani et al., 1989).

For the two subjects for whom EPG data were collected, the EPG frame corresponding to the mid-/t/ annotation point was used for analysis. The Q_p index was calculated for each of the 15 repetitions of /t/, separately for each vowel context and for each subject. The lists of Q_p values for mid-/t/ in the two vowel contexts were compared for significance, at the 0.05 level, separately for each subject, by means of two independent t-tests, using the Bonferroni adjustment (the cut-off value after the Bonferroni adjustment was 0.0025). If the difference was significant, it was concluded that there was a V-on-/t/ effect.

6.2.7. V-on-C coarticulation, analysis in Matlab

V-on-C coarticulation was measured by comparing ultrasound curves at the consonant closure for the consonant /t/, in the context of two vowels, /i/ and /a/ (abbreviated as "t_i curves" and "t_a curves", respectively; these and further abbreviations used in this chapter can be found in the section "Main abbreviations used in this text", at the beginning of the thesis). For each subject, the two sets of curves were compared, using the procedure based on the Nearest Neighbour technique (see Section 3.4.6 for details of the technique). The procedure described in Section 3.4.7.2 was used: comparing one set of across-group distances to two sets of within-group distances. A Univariate ANOVA was conducted in SPSS for each subject separately. The Games-Howell Post Hoc test was used to check for significant differences. If the test showed significant differences between across-group variation and both within-group variations, at the 0.05 level, then the distance between the two sets of curves was considered significant, and it was concluded that there was a V-on-C effect.

6.2.8. C-on-V coarticulation, analysis in Matlab

In order to measure C-on-V coarticulation, the tongue curves for /a/ were compared in the context of two consonants, /t/ and /k/. For each subject, the 15 $a1_t$ curves were compared with the 15 $a1_k$ curves, and the 15 $a2_t$ curves were compared with the 15 $a2_k$ curves (Table 6-1 clarifies the meanings of these abbreviations; see also the section "Main abbreviations used in this text" at the beginning of the thesis). Comparisons for significant difference were made using the same procedure as the one employed in the V-on-C coarticulation analysis described in Section 6.2.7. A Univariate ANOVA was conducted in SPSS for each subject separately. The Tukey HSD test was used when the assumption of equal variances of the dependent variable was not violated. The Games-Howell Post Hoc test was used when the assumption of equal variances of the dependent variable was violated. If the Post Hoc test showed significant differences between across-group variation and both within-group variations for $a1_k$ curves and $a1_t$ curves curves, at the 0.05 level, then the distance between the two sets of curves was considered significant, and it was concluded that there was a C-on-V1 effect. If the Post Hoc test showed significant differences between across-group variation and both within-group variations for $a2_k$ curves and $a2_t$ curves curves, at the 0.05 level, then the distance between the two sets of curves was considered significant, and it was concluded that there was a C-on-V2 effect.

	V1	V2
ata	$a1_t$	$a2_t$
aka	$a1_k$	$a2_k$

Table 6-1. Abbreviations used for sets of /a/ vowel curves, in analysis of C-on-V coarticulation. V1 curves from /ata/ sequences are called $a1_t$ curves; V1 curves from /aka/ sequences are called $a1_k$ curves; V2 curves from /ata/ sequences are called $a2_t$ curves; V2 curves from /aka/ sequences are called $a2_k$ curves.

6.2.9. Comparison of V-on-C and C-on-V coarticulation, analysis in Matlab

For measuring the difference between V-on-C and C-on-V effect, the sizes of these effects were calculated and compared. The size of the effect was determined by calculating average nearest neighbour distances between two sets of curves representing the same sound in two different contexts (see the description of this procedure in Section 3.4.7.1). The size of the V-on-C effect was indicated by the average distance between two sets of /t/ curves in the two different vowel contexts (t_i curves and t_a curves). The size of the C-on-V effect was calculated separately for V1 and V2. The size of C-on-V1 effect was indicated by the average distance between two sets of /a1/ curves in the two consonant environments ($a1_k$ curves and $a1_t$ curves). The size of C-on-V2 effect was indicated by the average distance between two sets of /a2/ curves in the two consonant environments ($a2_k$ curves and $a2_t$ curves). Statistical comparison was made by means of a Univariate ANOVA: the sets of average nearest neighbour distances were compared for significance (for further details, see Section 3.4.7.1). An ANOVA was conducted in SPSS for each subject separately. The V-on-C effect was compared with the C-on-V1 and C-on-V2 effects. The Games-Howell Post Hoc test was used to check for significant differences. If the test showed significant differences between the V-on-C effect and both C-on-V effects, then it was concluded that the V-on-C effect was significantly greater than the C-on-V effect.

6.2.10. Comparison of V1-C and C-V2 distances, analysis in Matlab

For calculating the difference between the V1-on-C and V2-on-C effect, the procedure described in Section 3.4.7.1 was used. The size of the effect was determined by calculating average nearest neighbour distances, for the three experimental stimuli (/aka/, /ata/ and /iti/), between V1 and C, and between C and V2. The size of the V1-on-C effect was indicated by the average distance between the set of V1 curves and the set of C curves. The size of the V2-on-C effect was indicated by the average distance between the set of V2 curves and the set of C curves. Statistical comparison was made in SPSS. A Univariate ANOVA was conducted separately for each subject. Average nearest neighbour distances were compared for V1-C versus V2-C (the independent variable was called syllable affiliation of the vowel), and for the three VCV types, /aka/, /ata/ and /iti/ (the independent variable was called VCV type). Pairwise comparison with the Bonferroni adjustment was used to look for significant differences depending on the syllable affiliation of the vowel. If the C-V2 distance was significantly smaller than the V1-C distance, at the 0.05 level, it was concluded that there was a syllable boundary effect on the coarticulation in the VCV sequence. The Games-Howell Post Hoc test was used to check for significant differences depending on the VCV type. If differences occurred, at the 0.05 level, it was concluded that V1-C and C-V2 distances in VCV sequences depended on the vowel environment.

6.3. Results

6.3.1. Overview

Average ultrasound tongue curves, taken at the annotation points described in Section 6.2.4, for the three different stimuli (/aka/, /ata/, /iti/) in the three subjects are presented in Figures 6-3 – 6-5. The figures demonstrate individual differences in tongue shapes. Some of the variation is due to the orientation of the ultrasound transducer under the chin.

A noticeable difference between these data and the tongue contour data presented in Chapters 4 and 5 is that in these data the tongue changes its shape much more throughout the VCV sequences than was the case with the labial consonants and /h/.

Another observation we can make from looking at the average tongue contours in the three subjects, is that the V2 curve generally appears to be closer to the C curve than the V1 curve.

Below, the three VCV sequences will be described separately.

Experiment 2: British English /h/

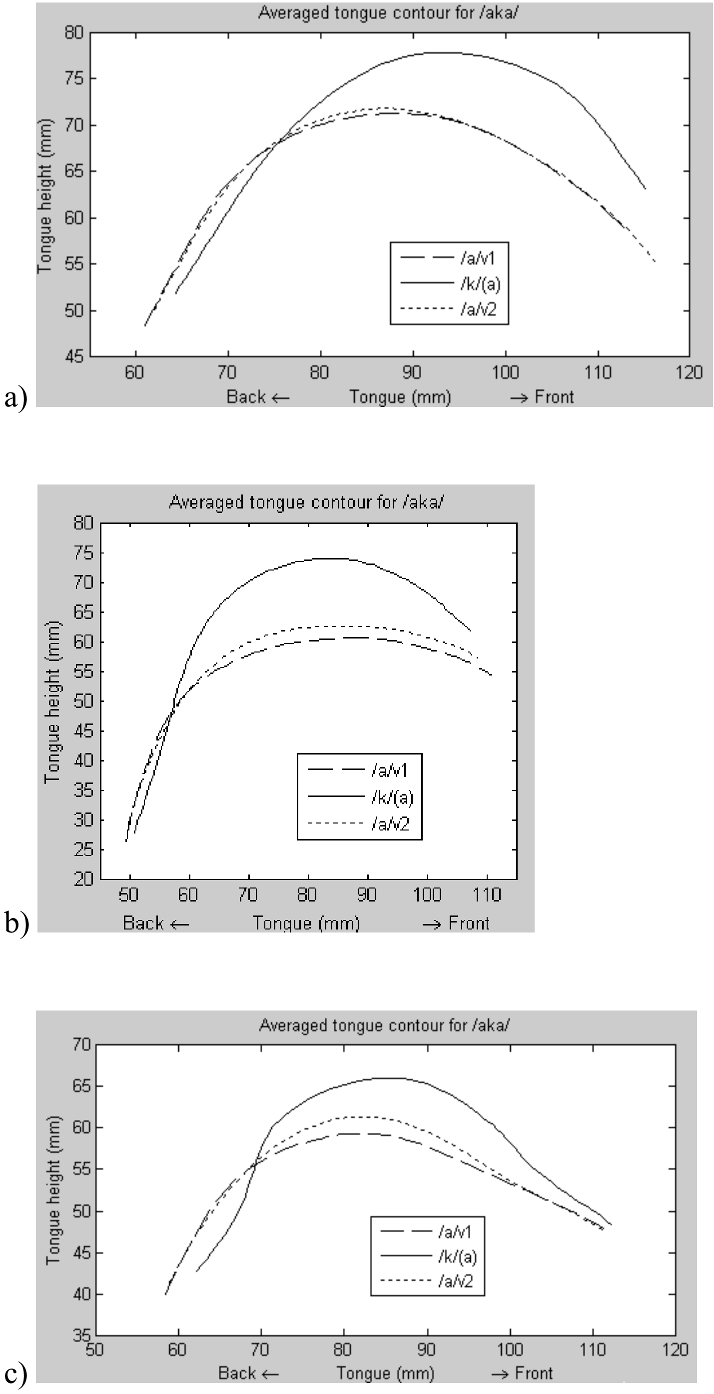


Figure 6-3. Average tongue contours in /aka/, in the three subjects: a) S1; b) S2; c) S3. Dashed line – V1 curve; solid line – C curve; dotted line – V2 curve.

In all the subjects, in /aka/ sequences (Figure 6-3), the tongue dorsum is noticeably higher in the consonant than in the surrounding vowels. This raised position of the dorsum is required for the occlusion formation. The blade is also higher in the consonant than in both vowels. The back part of the tongue is further forward in the consonant than in both vowels.

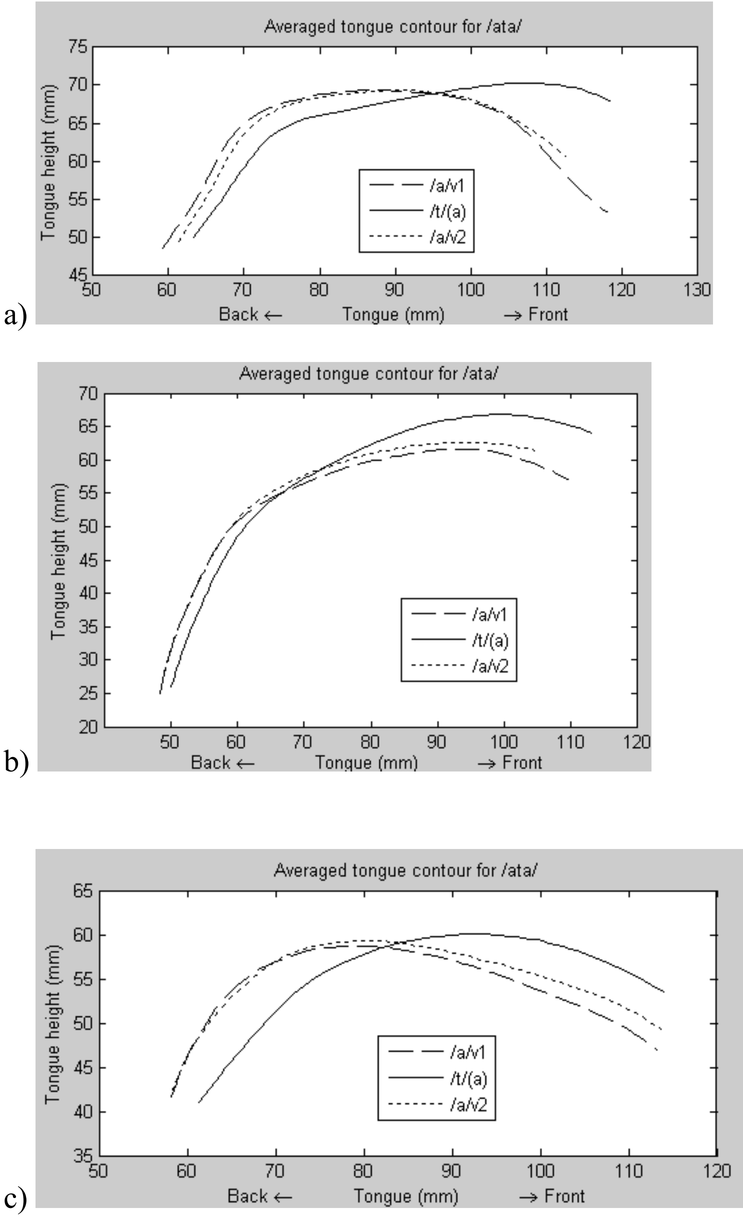
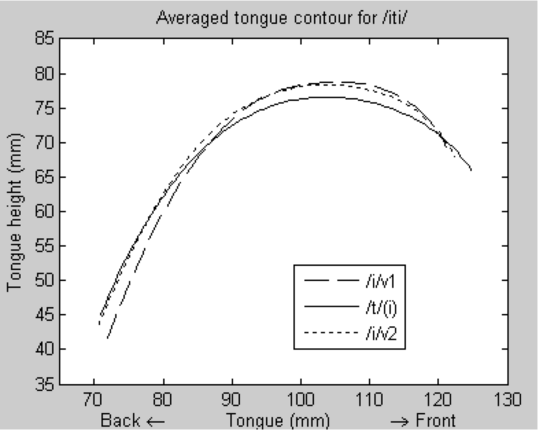


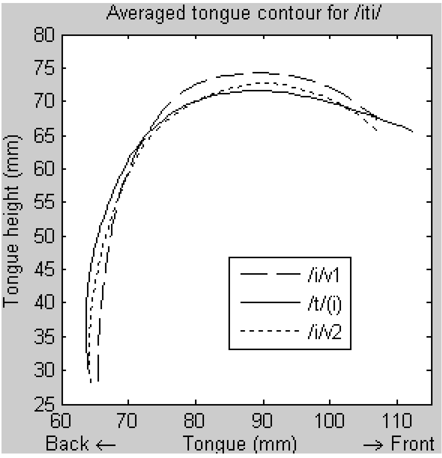
Figure 6-4. Average tongue contours in /ata/, in the three subjects: a) S1; b) S2; c) S3. Dashed line – V1 curve; solid line – C curve; dotted line – V2 curve.

In /ata/ sequences (Figure 6-4), the front part of the tongue is involved in making a consonant closure, and we can see that in all three subjects. The front third to a half of the tongue contour in the consonant is higher than both vowel curves. The back part of the tongue in the consonant is fronted in relation to both vowels (to a different degree in different subjects).

a)



b)



c)

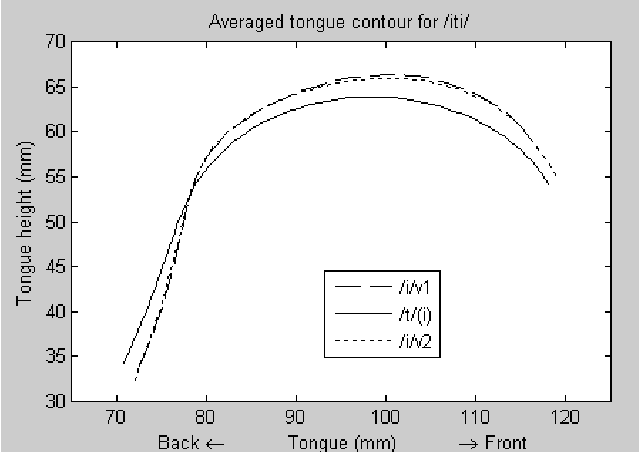


Figure 6-5. Average tongue contours in /iti/, in the three subjects: a) S1; b) S2; c) S3. Dashed line – V1 curve; solid line – C curve; dotted line – V2 curve.

In /iti/ (Figure 6-5), in all the subjects there is a tendency for the middle part of the tongue to be lower in /t/ than in both surrounding vowels, and for the tongue root to be more fronted during both vowels. We also notice that the consonant curve is closer to the vowels' curves in /iti/ sequences than in the sequences with the vowel /a/. This issue will be further explored in Section 6.3.6, where some quantitative data will be presented.

6.3.2. V-on-C coarticulation, EPG results

V-on-C coarticulation was measured using EPG, by comparing the contact pattern for /t/ in the context of two vowels, /i/ and /a/. EPG contact patterns over 15 repetitions are presented in Figure 6-6 for two subjects.

Experiment 2: British English /h/

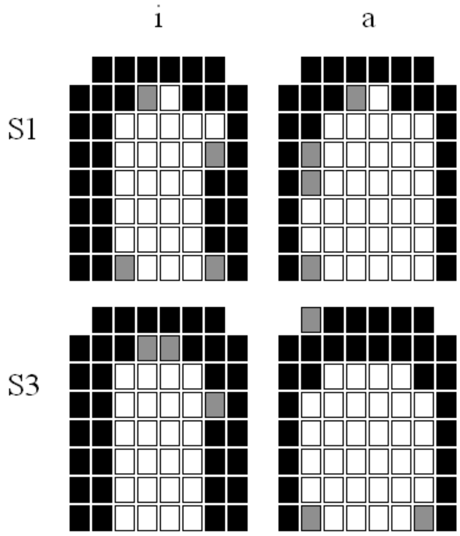


Figure 6-6. EPG contact patterns over 15 repetitions of /t/ in the context of /i/ (on the left) and in the context of /a/ (on the right), in two subjects. Black: 80-100% electrode activation; grey: 40-80% electrode activation; white: less than 40% electrode activation.

The values of the Q_p index (percentage of electrode activation in the palatal zone, i.e., the three back rows of the palate) are given in Table 6-2.

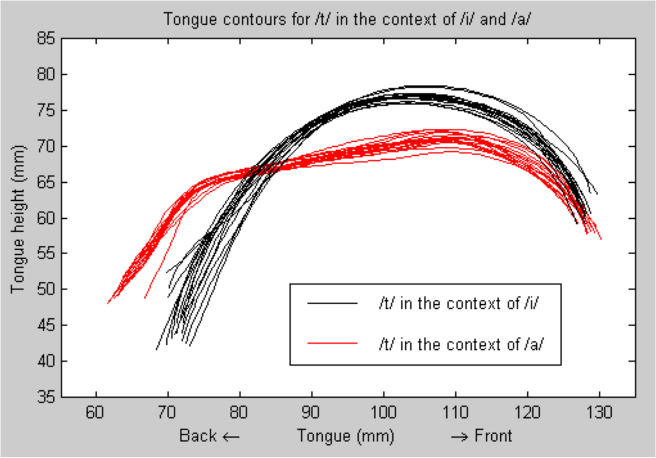
	i	a
S1	53%	29%
S3	51%	31%

Table 6-2. The Q_p index values over 15 repetitions of /t/ in the context of /i/ and in the context of /a/, in two subjects.

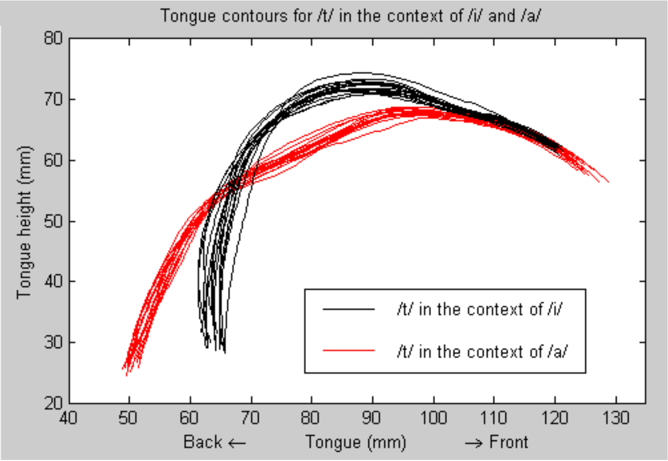
Statistical calculations produce highly significant results ($p < 0.001$): the EPG contact pattern is significantly different in the context of these two vowels, and there is a V-on-C effect.

6.3.3. V-on-C coarticulation, ultrasound results

In Figure 6-7, 15 t_a curves and 15 t_i curves are presented. In all the subjects, the only part of the ultrasound curves in which there is appreciable overlap between the two sets, is in the front part of the tongue, its length varying among the subjects. The rest of the tongue curve is noticeably different in the two vowel environments, resembling the neighbouring vowel contour (see Figures 6-3 – 6-5 for comparison). Note that there are some inter-speaker differences in tongue contours due to individual tongue shapes. For example, in S1, the back third of the tongue contour for the t_a curves is at a fairly sharp angle to the rest of the tongue, while in S3, the set of t_a curves is shaped as a semicircle, without such a noticeable change. Some other cross-subject differences are related to the transducer orientation under the chin. For example, in S2, the transducer was rotated anti-clockwise, as compared with the other two subjects. As a consequence of this, a longer stretch of the tongue blade was imaged in this subject than in the other two subjects. Hence, we can see that the overlapping part of the tongue contour for this subject is greater than for the other subjects.



a)



b)

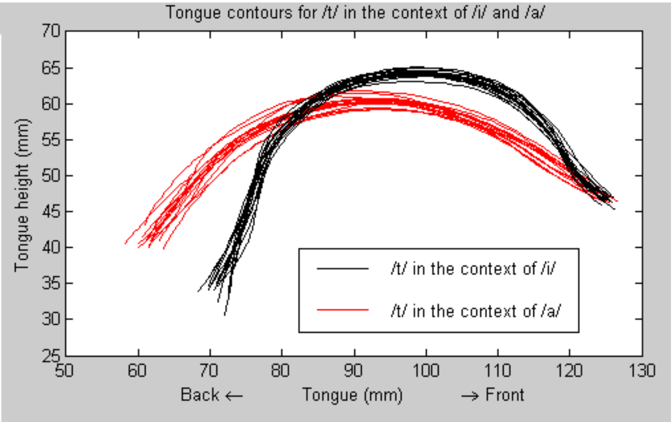
c) 

Figure 6-7. Tongue contours for 15 repetitions of /t/ in two vowel environments in the three subjects: a) S1; b) S2; c) S3.

Average distances between sets of curves are presented in Table 6-3.

	Across-group, $t_a - t_i$	Within-group, t_a	Within-group, t_i
S1	4.72	0.85	1.15
S2	4.70	0.85	1.18
S3	4.47	0.72	0.89

Table 6-3. Average distances between sets of curves, in millimetres, for t_a curves and t_i curves.

A Univariate ANOVA was conducted for the three subjects separately: across-group distances for t_i curves and t_a curves were compared with within-group distances for these curves (see Section 3.4.7.2 for the description of the procedure). In all the subjects, there was a significant effect: in S1, $F = 2778.33$, $df = 2$, $p \leq 0.001$; in S2, $F = 1944.28$, $df = 2$, $p \leq 0.001$; in S3, $F = 3390.27$, $df = 2$, $p \leq 0.001$. The Games-Howell Post Hoc test demonstrates that across group distances were significantly greater than within group distances, at the 0.05 level for all the subjects. These results show that in all the subjects there is a significant V-on-C effect.

In Figure 6-8, there are ultrasound curves for the two consonants, /k/ and /t/, taken from /aka/ and /ata/ sequences, respectively. We can easily see that in all the subjects, the forward part of the tongue in /k/ is lower than in /t/. The tongue dorsum is considerably higher in /k/ across all subjects. The tongue root position overlaps considerably in the two consonants’ curves, reflecting the vowel’s influence.

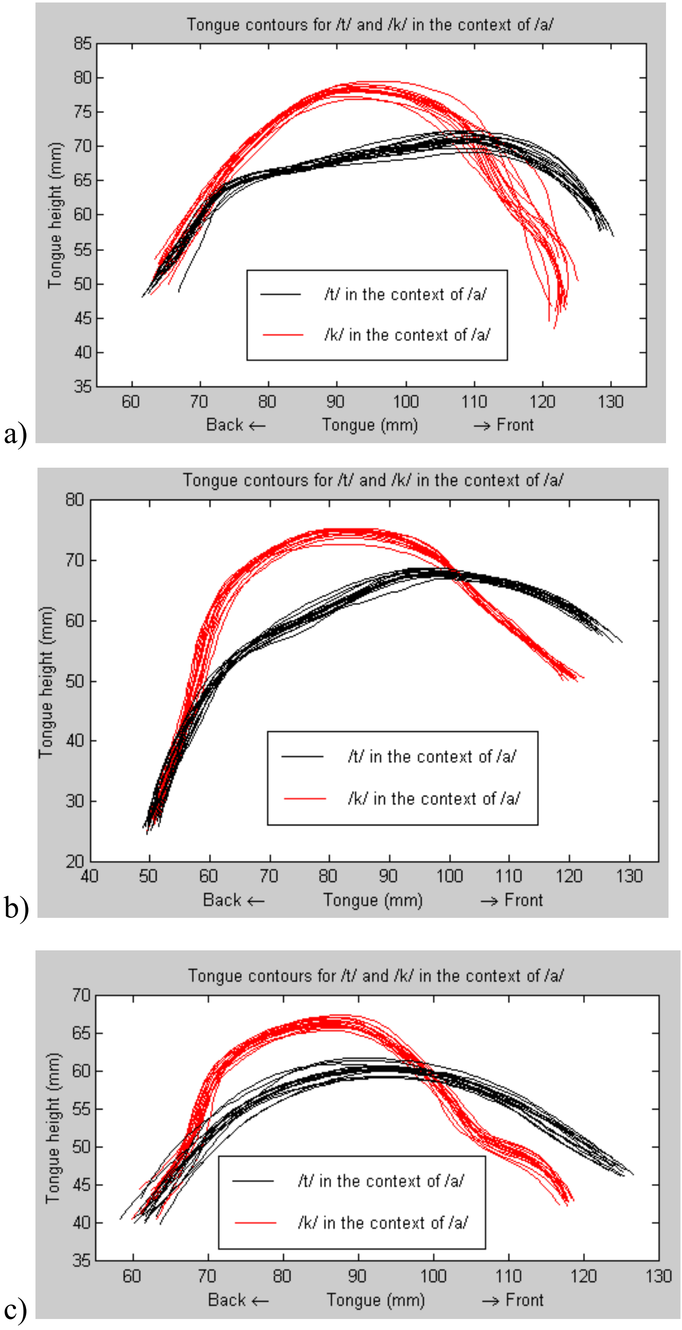
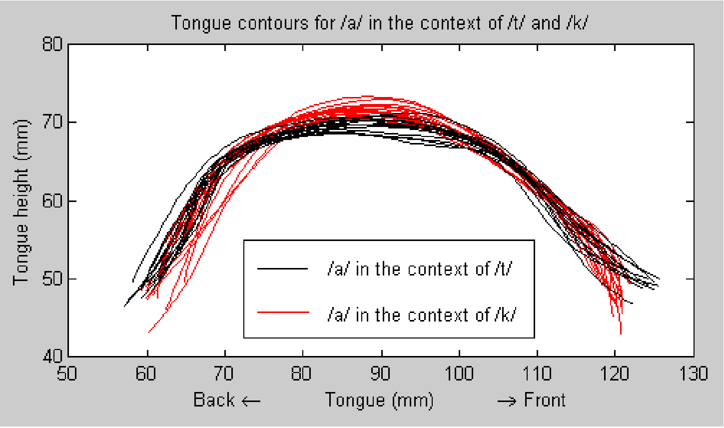


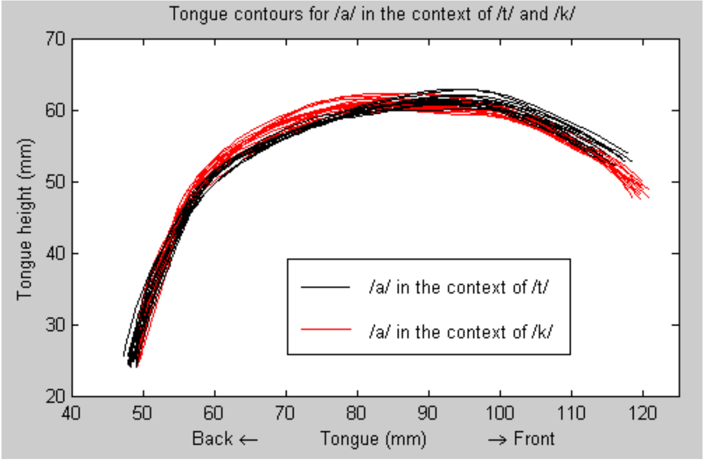
Figure 6-8. Tongue contours for 15 repetitions of /t/ and /k/ in the context of /a/ in the three subjects: a) S1; b) S2; c) S3.

6.3.4. C-on-V coarticulation

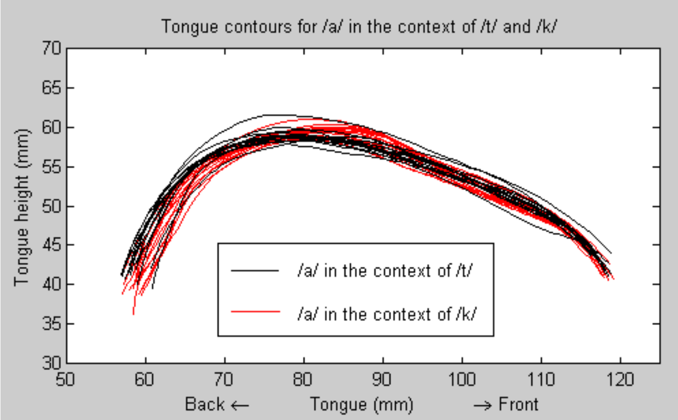
In Figures 6-9 and 6-10, tongue curves for the three subjects, for 15 repetitions of /a/ in two consonant contexts, /k/ and /t/, are presented: V1 and V2 separately. The tongue root appears to be very similar in the two consonant contexts. The difference in the tongue dorsum position is obvious in the two consonant contexts: the tongue dorsum is higher in the /k/ context in all the subjects. The front part of the tongue is generally slightly lower in the /k/ context. These differences are less clear in the V1 in subject S3, where the V1 contours are more comparable in the two consonant contexts than the V2 contours (compare Figure 6-9c and Figure 6-10c).



a)

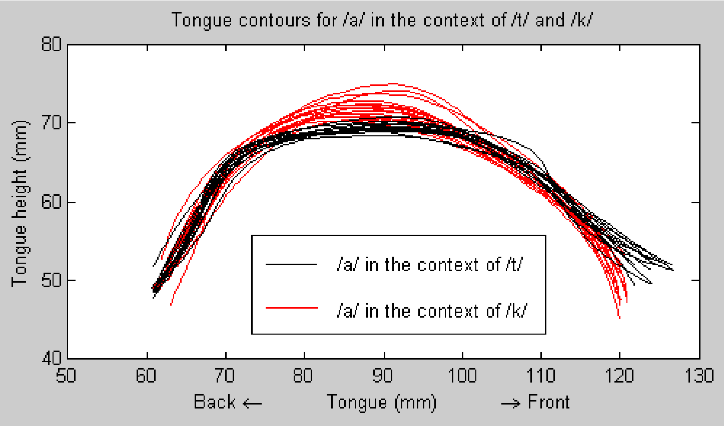


b)

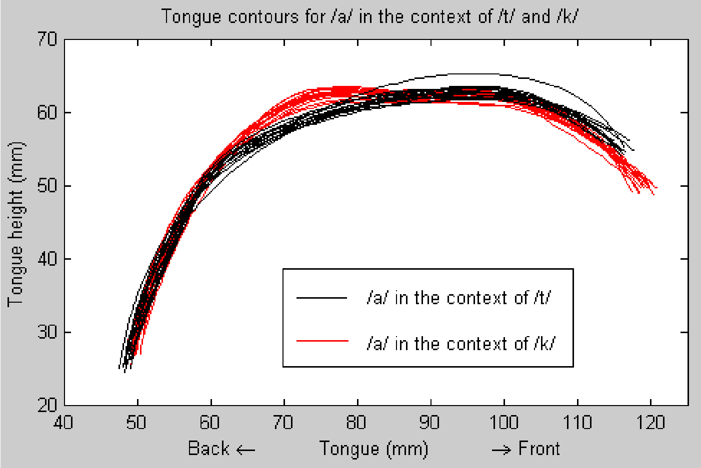


c)

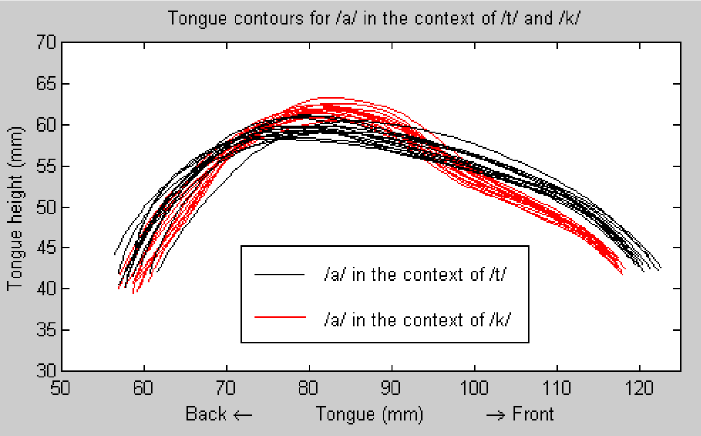
Figure 6-9. Tongue contours for 15 repetitions of V1 (/a/) in two consonant environments in the three subjects: a) S1; b) S2; c) S3.



a)



b)



c)

Figure 6-10. Tongue contours for 15 repetitions of V2 (/a/) in two consonant environments in the three subjects: a) S1; b) S2; c) S3.

Average distances between sets of curves are presented in Table 6-4.

	V1			V2		
	Across-group, $a1_k - a1_t$	Within-group, $a1_k$	Within-group, $a1_t$	Across-group, $a2_k - a2_t$	Within-group, $a2_k$	Within-group, $a2_t$
S1	1.47	1.11	1.08	1.52	1.15	0.84
S2	1.17	0.79	0.72	1.33	0.68	0.85
S3	1.13	0.95	0.95	1.91	0.93	1.13

Table 6-4. Average distances between sets of curves, in millimetres, for a_k curves and a_t curves.

First, a Univariate ANOVA was conducted for V1 in the two consonant environments, for the three subjects separately: across-group distances for a_k curves and a_t curves were compared with within-group distances for these curves (see Section 3.4.7.2 for the description of the procedure). In all the subjects, there was a significant effect: in S1, $F = 42.33$, $df = 2$, $p \leq 0.001$; in S2, $F = 181.85$, $df = 2$, $p \leq 0.001$; in S3, $F = 11.80$, $df = 2$, $p \leq 0.001$. The Games-Howell Post Hoc test conducted for S1 and S3 demonstrates that across group distances were significantly greater than within group distances, at the 0.05 level. The Tukey HSD Post Hoc test conducted for S2 shows that for this subject, across group distances were also significantly greater than within group distances, at the 0.05 level. These results show that in all the subjects there is a significant C-on-V1 effect.

After that, a Univariate ANOVA was conducted for the second vowel in the two consonant environments, for the three subjects separately. In all the subjects, there was a significant effect: in S1, $F = 134.63$, $df = 2$, $p \leq 0.001$; in S2, $F = 229.89$, $df = 2$, $p \leq 0.001$; in S3, $F = 266.14$, $df = 2$, $p \leq 0.001$. The Games-Howell Post Hoc test demonstrates that across group distances were significantly greater than within group distances, at the 0.05 level for all the subjects. These results show that in all the subjects there is a significant C-on-V2 effect.

6.3.5. Comparison of V-on-C and C-on-V coarticulation

In Table 6-5, average values of nearest neighbour distances and standard deviations are given for all the subjects, representing V-on-C and C-on-V coarticulatory effects. The V-on-C effect is represented by the distance between the 15 t_i curves and the 15 t_a curves. The C-on-V1 effect is represented by the distance between the 15 $a1_t$ curves and the 15 $a1_k$ curves. The C-on-V2 effect is represented by the distance between the 15 $a2_t$ curves and the 15 $a2_k$ curves.

Subject	Distance $t_i – t_a$ (V-on-C effect)		Distance $a1_t – a1_k$ (C-on-V1 effect)		Distance $a2_t – a2_k$ (C-on-V2 effect)	
	Mean	SD	Mean	SD	Mean	SD
S1	4.72	0.64	1.47	0.46	1.52	0.39
S2	4.70	0.78	1.17	0.25	1.33	0.25
S3	4.78	0.55	1.13	0.37	1.91	0.38

Table 6-5. Average values of nearest neighbour distances and standard deviations for the three subjects, in millimetres. The distance $t_i – t_a$ represents the V-on-C effect; the distance $a1_t – a1_k$ represents the C-on-V1 effect; the distance $a2_t – a2_k$ represents the C-on-V2 effect.

The table demonstrates that the distances between consonant curves in different vowel environments are greater than the distances between vowel curves in different consonant contexts.

A Univariate ANOVA was conducted, for the three subjects separately, with the aim to compare the three distances. In all the subjects, there was a significant effect: in

S1, $F = 3029.27$, $df = 2$, $p \leq 0.001$; in S2, $F = 3557.43$, $df = 2$, $p \leq 0.001$; in S3, $F = 4289.32$, $df = 2$, $p \leq 0.001$. The Games-Howell Post Hoc test conducted for S1 demonstrates a significant difference (at the 0.05 level) between the $t_i - t_a$ distance and both $a_t - a_k$ distances, and no significant difference between $a1_t - a1_k$ and $a2_t - a2_k$ distances. The Games-Howell Post Hoc test results for S2 and S3 demonstrate a significant difference (at the 0.05 level) between the $t_i - t_a$ distance and both $a_t - a_k$ distances, and also that the $a2_t - a2_k$ distance is significantly greater than the $a1_t - a1_k$ distance, at the 0.05 level.

These results demonstrate that the influence of vowels on consonants is greater than the influence of consonants on vowels.

6.3.6. Comparison of V1-C and C-V2 distances

In Table 6-6, average distances between V1 and C curves and between C and V2 curves are presented for all the subjects and for all three stimuli: /aka/, /ata/, and /iti/. It is clear from the table that in all the subjects the distance between the C curve and the V2 curve is smaller than the distance between the V1 curve and the C curve. A Univariate ANOVA was conducted, separately for each subject, to explore how different these two distances are, and also whether they depend on the VCV type. The results for all three subjects show that there was a significant effect of the syllable affiliation of the vowel (V1 or V2) on the consonant-vowel distances (S1: $F = 300.82$, $df = 1$, $p \leq 0.001$; S2: $F = 783.13$, $df = 1$, $p \leq 0.001$; S3: $F = 341.81$, $df = 1$, $p \leq 0.001$). Pairwise comparison was made, using the Bonferroni adjustment, and its results show that the C-V2 distance was significantly smaller than the V1-C distance, at the 0.05 level.

		aka	ata	iti
S1	V1-C	4.27	5.26	1.97
	C-V2	3.95	4.43	1.57
S2	V1-C	6.17	3.99	2.51
	C-V2	5.02	3.21	2.00
S3	V1-C	3.55	4.48	1.59
	C-V2	2.89	3.59	1.53

Table 6-6. Average distances in millimetres between V1 and C and between C and V2, for the three subjects.

The ANOVA results also feature a significant effect of VCV type on the distances (S1: $F = 3815.55$, $df = 2$, $p \leq 0.001$; S2: $F = 4519.56$, $df = 2$, $p \leq 0.001$; S3: $F = 2533.20$, $df = 2$, $p \leq 0.001$). The Games-Howell Post Hoc test demonstrates significant differences between all pairs of VCVs (/aka/ versus /ata/, /ata/ versus /iti/ and /aka/ versus /iti/), for all the subjects, at the 0.05 level. We can see in the table that in the /i/ context, consonant-vowel distances are much smaller than in the other two vowel contexts. This difference between /i/ and /a/ contexts will be discussed in Section 6.4.7.

A significant interaction was observed in all the subjects, between the two independent variables of VCV type (/aka/, /ata/ or /iti/) and syllable affiliation of the vowel (V1 or V2): in S1, $F = 27.70$, $df = 2$, $p \leq 0.001$; in S2, $F = 40.69$, $df = 2$, $p \leq 0.001$; in S3, $F = 73.56$, $df = 2$, $p \leq 0.001$.

6.4. Discussion

6.4.1. V-on-C and C-on-V coarticulation

The EPG experiment produced results quite consistent with those presented in Recasens et al. (1997). A significant difference was demonstrated between the contact pattern of /t/ in the context of /i/ versus /a/. These results support Hypothesis 1.

The results of analysing ultrasound data in this experiment were consistent with the EPG results. It was shown, across subjects, that the consonant /t/ had significantly different tongue shapes when it occurred in two different vowel contexts, /i/ versus /a/. These findings support Hypothesis 2. The tongue root was more retracted in the /a/ context than in the /i/ context. The dorsum was lower in the /a/ context than in the /i/ context. The only part in the two /t/ contours that overlapped was the front part of the tongue. This can be explained by the fact that only the front part of the tongue was involved in making the occlusion for the alveolar consonant production.

In the DAC model terms, both EPG and ultrasound results signify that the degree of resistance to vocalic coarticulation in the British English consonant /t/ in VCV sequences is lower than maximal, i.e., that there is a significant V-on-/t/ effect. We can also confirm that the front part of the tongue, where the overlap was observed in the two vowel contexts, is more resistant to vocalic coarticulation than the rest of the tongue.

More ultrasound evidence of vocalic influence on consonant production was provided in Section 6.3.3, where the tongue root position was shown to overlap considerably in /k/ and /t/ in the context of /a/. This overlap can be interpreted as an influence of the vowel /a/, which has a retracted tongue root.

In terms of the V-on-C effect, there is the following difference between the results of this experiment and the first two experiments. In non-lingual consonants (bilabial stops and /h/), there was no overlap in the tongue curves for the consonant in the three vowel contexts, because for these consonants there was no particular requirement on the tongue, common for the three vowel environments. So the three experiments produced evidence that the lingual shape in bilabial consonants and /h/ is more influenced by the neighbouring vowels in VCV sequences than in lingual consonants. Within the DAC framework, we can say that the degree of resistance to lingual coarticulation is greater in lingual consonants than in non-lingual consonants.

A significant C-on-V effect was demonstrated in the ultrasound data in Experiment 3. It was shown, across subjects, that the vowel /a/ had significantly different tongue shapes when it occurred in two different consonant contexts, /k/ versus /t/. These results support Hypothesis 3. The tongue root position was comparable in the two different consonant contexts, the tongue dorsum was slightly higher in the /k/ context than in the /t/ context, and the blade was slightly lower in the /k/ context than in the /t/ context. In the DAC model terms, the results of this experiment mean that the degree of resistance to consonantal influence in the British English vowel /a/ in VCV sequences is lower than maximal, i.e., that there is a significant C-on-/a/ effect. The results also suggest that the degree of CR of the tongue root to consonantal influence is higher than that of the rest of the tongue, because there was more overlap in the root position than in the dorsum and the blade.

The focus of Experiment 3 was to some extent methodological. For this reason, the data set was kept somewhat restricted. A comprehensive data set would have included /iki/ sequences.

The influence of different vowels on /t/ was demonstrated to be significantly greater than the influence of different consonants on /a/. Specifically, the difference between the vowel contours in the two different consonant contexts was significantly smaller than the difference between the consonant contours in the two vowel contexts. This result supported Hypothesis 4. The DAC model interpretation is that the vowel is more resistant than the consonant to the neighbouring segments in VCV sequences.

Some implications of this result for the CR theory are that the distribution of resistance along the tongue contour appears to be different in consonants and vowels. In the consonant /t/, the blade had a similar position in different vowel contexts, and the rest of the tongue varied greatly, according to the surrounding vowels. In the vowel /a/, while the position of the tongue root was similar across consonant contexts, the rest of the tongue had a much more similar position in two different consonantal contexts than the rest of the tongue in the consonant /t/ in two different vocalic contexts. This suggests that, while both in consonants and in vowels, one part of the tongue appears to be more resistant to coarticulation than others, there is a difference in the degree of resistance of the remaining part of the tongue contour. According to our data, this remaining part of the tongue contour is much less resistant in consonants than in vowels. In Chapter 7, this observation will be further pursued, when quantifying CR of consonants and vowels.

6.4.2. Syllable/word boundary influence on VCV coarticulation

It has been demonstrated in the literature that a syllable boundary greatly affects coarticulatory patterns in V#CV sequences (e.g., Kozhevnikov and Chistovich 1965; Gay 1977; Perkell 1986; Browman and Goldstein 1988; Byrd 1995; Lindblom et al. 2002; see also Section 2.3.3). In all three experiments in this work, the data were VCV sequences with the syllable boundary after the first vowel. In Experiment 1 (Chapter 4), Russian nonsense VCV sequences presented to the subjects were each spelt in one word. In Russian words having VCV structure, syllable boundary is always between the first vowel and the consonant. So in this experiment, the consonant belonged to the second syllable. In Experiment 2 (Chapter 5), the data were English VhV sequences. According to the rules of English phonotactics, /h/ cannot be syllable-final. So in this case, we also had the consonant belonging to the second syllable.

In Experiment 3, the second vowel of the VCV was shown to influence the consonant tongue configuration more than the first vowel, across subjects. It was demonstrated that the distance between the C curve and the V2 curve was significantly smaller than the distance between the V1 curve and the C curve. As argued in the previous paragraph, it seems plausible to attribute this to the presence of a syllable boundary after the first vowel. However, in this third experiment, there was a word boundary at this point as well. This raises the possibility that the differential effect on the second vowel was (partly, at least) due to the presence of the word boundary rather than the syllable boundary. I have, of course, no data from the experiment which can be used to distinguish between a possible syllable boundary effect and a possible word boundary effect, and this should be borne in mind when reading the rest of this section.

In the first experiment, with Russian bilabial consonants, qualitative observations were made of average three-curve graphs over 15 repetitions of the same VCV. According to these observations, the V2 curve appeared to be closer to the C curve than the V1 curve was. This pattern was consistent across subjects and across VCV types. This fact was interpreted as the influence of syllable boundary on segmental coarticulation: the consonant was coarticulated more strongly with the second vowel, i.e., with the vowel with which it formed the same syllable. In Experiment 2, when analysing British English VhV sequences, in addition to qualitative observations, the effect was measured using a quantitative procedure designed after the first experiment. “Eye-balling” the three-curve graphs in Experiment 2 produced the impression that the consonant curve was often closer to the second vowel than to the first vowel, in many tokens. The quantitative procedure was applied to the data from two subjects, and the results showed that in one subject the C-V2 distance was significantly smaller than the V1-C distance, and in the other subject it was not. So only partial support was obtained for stronger coarticulation in the CV-complex than in the VC-complex in VhV sequences.

Interpreting the results of Experiments 1 and 3 in terms of the DAC model, we can say that position in relation to the syllable/word boundary influences the degree of CR exhibited by the consonant: the consonant is less resistant to the vowel belonging to the same syllable than to the vowel belonging to a different syllable. Possible reasons for the lack of uniformity between subjects in VhV sequences (Experiment 2) were presented in Section 5.4.6, when discussing the results of Experiment 2. Very small distances between the ultrasound curves for V1, C and V2 were suggested to be a possible reason. In Experiment 1 (Chapter 4), there were substantial differences between the V1 curve and the V2 curve, explained by the stress influence. In Experiment 3 (Chapter 6), there were differences between the C curve and the two vowel curves, due to conflicting articulatory requirements on production of the consonant and the flanking vowels. In Experiment 2, in VhV sequences, the subject who did not exhibit the syllable boundary influence on VCV coarticulation had significantly smaller V1-C and C-V2 distances than the subject who did exhibit the syllable boundary effect. These small distances could have been a factor limiting the extent of the syllable boundary influence on VCV coarticulation. Another factor contributing to the cross-subject difference could have been as follows. The subject who did not exhibit the syllable boundary influence produced both syllables of the VhV sequence with equal stress. The other subject, who did have a syllable boundary effect, realised the VhV sequences with the second syllable stressed. Stress affected the vowel quality in that subject (the first vowel was realised as a schwa, and the second vowel was produced as [ɑ]), and it could have induced the pattern where C-V2 distance was closer than the V1-C distance. The following interpretation could be suggested within the DAC framework, based on the data from Experiment 2. Possibly in the consonants that have a low degree of resistance to lingual coarticulation of the surrounding vowels (like /h/), and are greatly coarticulated with them, there may be no such a big sensitivity to the syllable boundary as in the consonants that are generally more resistant to the surrounding vowels, such as lingual consonants. And manifestation of syllable boundary influence on segmental coarticulation of non-lingual consonants may depend on individual subjects more than in the case of lingual consonants. However, these suggestions would need more experimental support, because of the observed cross-subject variation in Experiment 2 and the presence of the word boundary in Experiment 3.

One question that arises from the discussion in this section is about the relevance of taking details of individual variation into account when explaining experimental data. On the one hand, linguistic studies aim to find general patterns that are common across speakers. On the other hand, it is important to discuss various factors that induce cross-subject variability. Knowing the limits of articulatory variability in speakers without speech disorders helps to establish the range of variation acceptable for successful communication, and to be more confident in studying populations that have disordered speech. A further question relating to individual variation is how we can identify a difference between tongue movements that occur because they are planned or programmed in terms of speech motor control and movements caused by purely biomechanical factors. This question is very challenging. It forms a part of a more general question: which changes in acoustic signal, auditory output and articulator movements are linguistically relevant, i.e., conveying information to the listener, and which changes are not relevant to the listener, i.e., form “noise”. A possible way to approach this question is not only to look for patterns in the speech signal, but to look for a linguistically relevant function that unites those patterns. An example from this experiment would be the following pattern found across subjects: advanced tongue root for /t/ in the context of /i/ and retracted tongue root for /t/ in the context of /a/. The function of this pattern could be to convey the information about the influence of two different contexts on the consonant, and hence to facilitate perception of this speech sound in both contexts.

6.4.3. Gestural compatibility and resistance to coarticulation

It was noted in Section 6.3.1 that the C curve appeared to be closer to the vowel curves in /iti/ sequences than in /ata/ and /aka/ sequences (see Figures 6-3 – 6-5). In Section 6.3.6, these impressions were numerically confirmed: the V1-C and C-V2 distances were shown to be significantly smaller in /iti/ sequences than in /ata/ and /aka/ sequences, across subjects. The DAC model claims that the degree of coarticulation depends, among other factors, on gestural compatibility of adjacent sounds, i.e., on how similar or different the articulatory requirements are for producing the neighbouring sounds (e.g., Recasens et al. 1997). An example is the difference between coarticulatory patterns in /ata/ and /iti/ sequences observed in our data. For producing the English consonant /t/, the tongue blade is raised to the alveolar ridge. The rest of the tongue, as we have seen in this experiment, has some freedom for coarticulation with adjacent sounds. For producing the vowel /a/, the tongue root is retracted, and the rest of the tongue is in a rather flat and low posture. For the production of /i/, the dorsum is raised, the root is advanced, and the blade, following the dorsum, is in a higher position than during the vowel /a/. So in /iti/, the tongue blade has less distance to travel from its position during the vowel production towards the place of the consonant occlusion. In /ata/ sequences, the blade has a bigger route to travel to make an occlusion. Thus, the gestural compatibility of /t/ and /i/ is greater than that of /t/ and /a/. This results in the smaller tongue displacement in /iti/ sequences observed in this experiment.

As described in Section 2.1.2, Recasens et al. (1997) assign the highest possible value of 3 to the vowel /i/, based on the fact that the tongue dorsum is directly involved in the production of this vowel. The DAC value assigned to /a/ is lower (2), because the tongue dorsum, according to Recasens and his colleagues, is not directly involved in the production of /a/. The consonant /t/ in VCV sequences also has a DAC value of 2, as "the tongue dorsum is not directly involved in closure formation but is subject to coupling effects with the primary articulator" (Recasens et al. 1997, p. 545). In the DAC model, assigning the DAC values to the sounds is based on tongue dorsum behaviour. Thus some of their descriptions of the DAC values come from negative definitions (e.g., "tongue dorsum not involved"). These descriptions do not fully explain the observed patterns. To explain the lingual coarticulatory patterns in /ata/ and /iti/ sequences observed in this work, it would be helpful to include specifications of active tongue parts, as in the previous paragraph.

Experiment 2: British English /h/

More examples of gestural compatibility from our results include Russian VCV sequences with bilabial stops and high vowels (Chapter 4). The tongue dorsum is raised and fronted in palatalised consonants and in the vowel /i/, so there is no discontinuity in lingual coarticulation in /iCʲi/ sequences. The tongue dorsum is raised and the tongue root is displaced backwards in non-palatalised consonants and in the vowel /u/, so there is no discontinuity in coarticulation in /uCu/ sequences. Also, the /h/ and the /a/ in the British English /aha/ sequences exhibit gestural compatibility. As argued in Experiment 2 (Chapter 5), the tongue root is specified for a retracted position both in /h/ and in /a/, so no discontinuity is observed in lingual coarticulation in /aha/ sequences.

6.4.4. Separate specifications for different parts of the tongue

In the two previous chapters, some speculations were presented on whether parts of the tongue, and not necessarily the whole tongue, may have separate specifications (see Sections 4.4.10 and 5.4.9). Examples of such interpretation based on the results from Experiment 3 are presented below.

In /aka/ sequences (Figure 6-3), across subjects, the tongue dorsum was noticeably higher in the consonant than in the surrounding vowels. This raised position of the dorsum is explained by the requirement to produce the occlusion. The blade was higher in the consonant than in both vowels. The root of the tongue was further forward in the consonant than in both vowels. These displacements of the blade and the root seem to be due to the tongue dorsum raising to produce a velar closure. The blade was probably following the dorsum, rather than actively raising. The root, in the case of /k/, presumably also passively deformed, following the active dorsum raising for producing the closure.

In /ata/ and /iti/ sequences (Figure 6-4 and 6-5), the front part of the tongue was involved in making a consonant closure. In /ata/, the front part of the tongue contour in the consonant was higher than both vowel curves. The tongue root in the consonant was fronted in relation to both vowels. The middle part of the tongue in this case seemed to be following the blade, and, together with the tongue root, to be contributing to the generally more fronted tongue position for the consonant as compared with that of the vowels. In /iti/ sequences (Figure 6-5), there was a tendency for the middle part of the tongue to be lower in /t/ than in both surrounding vowels, and for the tongue root to be more fronted during both vowels. In /iti/, like in /ata/, the blade also actively moved towards the occlusion position, and the rest of the tongue seemed to have accommodated to the position of the blade required for the occlusion.

In cases such as the British English VhV sequences with high vowels (e.g., Section 5.4.2), it is very difficult to show which parts of the tongue are controlled separately, and which tongue parts are active during the /h/ production. First, because it is not clear what experimental data would show this. Secondly, ultrasound, even though it gives us important information on the movement of the whole tongue body, is a technique that does not allow for tracking individual flesh-points on the tongue (unlike, e.g., EMA). A combination of ultrasound and EMA could be a good technique for attempting to study the behaviour of individual parts of the tongue in relation to the whole tongue body. One more complication is that the tongue is a part of the complex articulatory system, and jaw and larynx movements, as well as movements of the velum, affect movements of the tongue.

One argument for pursuing the kind of research that would produce some numerical evidence on separate specification of the different parts of the tongue comes from clinical studies of disordered speech. Gibbon (1999) describes the so-called “undifferentiated lingual gestures”. Undifferentiated gestures are defined in Gibbon (1999), based on EPG data, as “EPG patterns that have, at maximum constriction during singleton lingual target consonants, anterior midsagittal contact occurring simultaneously with midsagittal posterior contact” (Gibbon 1999, p. 387). Gibbon (1999) refers to previous research, which shows that different parts of the tongue are independently controlled (e.g., Hardcastle 1976; Nguyen et al. 1996; Farnetani 1997). Parts of the tongue named in Gibbon (1999) include tongue tip/blade, tongue body, and the lateral components of the tongue (see also Gibbon 2006). Another argument comes from gesture-based models of speech production. In articulatory phonology, there are separate gestures for tongue body and tongue tip (e.g., Browman & Goldstein 1990). In the later version of the DAC model, speech sounds have different DAC values for tongue dorsum and tongue front (e.g., Recasens 2004). Functional divisions of the tongue have been proposed in the literature. For example, Recasens and Pallarès (2001) distinguish four parts, relevant for analysing EPG data of tongue-palate contact: 1) tip and blade, 2) predorsum, 3) mediodorsum, 4) postdorsum. Stone et al. (2004) reported some evidence that "tongue deformation is controlled by the synergistic coordination of 'functional segments' of the tongue" (Stone et al. 2004, p. 508).

The notion of independent control of different parts of the tongue is useful in phonetic research. By studying the regularities in displacement of tongue parts for achieving linguistically defined targets, a more detailed representation of speech motor control mechanisms can be made than by describing movements of the tongue as a whole. Identifying the range of variability for functionally driven movements of tongue parts in non-disordered speech production can also help to identify and investigate abnormal or delayed coarticulatory patterns, as seen in disordered speakers or very young children. The ultrasound-based methods developed in this work could be applied to studying motor control in parts of the tongue, in normal and abnormal populations, both adults and children.

6.5. Summary

In this experiment, resistance to coarticulation was measured in lingual stops and in vowels. Several hypotheses were formulated within the framework of the Degree of Articulatory Constraint model, and tested using ultrasound data.

EPG combined with ultrasound and acoustics was used for data collection and analysis. Similar results were obtained in the EPG experiment and in the ultrasound analysis of V-on-C coarticulation in lingual consonants. This was the first systematic experiment where midsagittal ultrasound data were used together with EPG data for measuring CR in speech.

Ultrasound and EPG results demonstrated a significant V-on-C effect. The EPG results were consistent with those presented in Recasens et al. (1997): the contact pattern for the consonant varied, depending on the vowel environment. Within the DAC model, the results were interpreted by claiming that the resistance of the intervocalic lingual consonant to the coarticulatory influence of the surrounding vowels is lower than maximal.

A significant C-on-V effect was reported, using ultrasound data. The results were interpreted within the DAC model by claiming that the resistance of the vowels to the coarticulatory influence of the intervocalic consonant is lower than maximal.

The V-on-C effect was significantly greater than the C-on-V effect. In terms of the DAC model, this finding was interpreted to suggest that resistance of vowels to the consonantal influence is greater than resistance of consonants to the vocalic influence.

Word boundary influence on VCV coarticulation was observed. This was interpreted as a manifestation by the intervocalic consonant of a smaller CR to the tautosyllabic V2 than to the V1.

The results of this experiment were discussed together with the results of the first two experiments, within the DAC framework. The discussion of specifications of tongue parts was continued. In Chapter 7, suggestions for quantifying CR of the sounds forming VCV sequences will be presented, with examples of calculations, based on the ultrasound data from this work.

7. PROPOSAL FOR A NEW METHOD OF QUANTIFYING COARTICULATION RESISTANCE

7.1. Introduction

In this chapter, the suitability of the Degree of Articulatory Constraint (DAC) model for explaining my experimental findings is explored. Implications are drawn for how we can better account for the mechanisms of speech production. Implications of the experimental results are presented for developing and enlarging the DAC model within the Coarticulation Resistance (CR) approach to speech production. The applicability of ultrasound as a research technique for studying coarticulation is discussed.

7.2. Critical evaluation of the DAC model in relation to this study

The results from ultrasound experiments in this work cannot be directly compared to most results of the EPG studies of VCV coarticulation within the DAC model. Differences in methodologies and their implications for interpreting ultrasound results within the DAC framework will be described in this section. Suggestions for enlarging the DAC model will be presented.

In this study, V-on-C coarticulation was measured by comparing tongue shapes for the same consonant in two different vowel environments. The question about V-on-C coarticulation was how different the consonant tongue shape is in two different vowel contexts. In Recasens et al. (1997) and in other studies using the same methodology, symmetrical and non-symmetrical vowel-consonant-vowel sequences were studied. The principal aim of the studies by Recasens and colleagues was to measure anticipatory and carryover coarticulation separately. Lingual contact patterns were compared for the consonant from, e.g., the pair of /aCa/ and /aCi/ sequences (for measuring anticipatory coarticulation) and, e.g., the pair of /aCa/ and /iCa/ sequences (for measuring carryover coarticulation), in order to answer the question to which degree changing the vowel environment affects the consonant. The midpoint of the consonant closure was used for the measurements.

C-on-V coarticulation in this study was measured using a similar procedure to the one used for measuring V-on-C coarticulation with ultrasound. Tongue shapes for the same vowel in two different consonant environments were compared. The question about C-on-V coarticulation was how different the vowel tongue shape is in two different consonant contexts. In Recasens et al. (1997) and in other studies using the same methodology, the data, like in the present work, were symmetrical VCV sequences. The criteria for defining the C-on-V effect, however, were different from the criteria used in this study. The researchers first identified a typical lingual contact pattern for the steady state of the vowel. Then they compared that pattern with lingual contact patterns in each EPG frame between the vowel steady state and the consonant. For calculating the C-on-V1 effect ("C-to-V anticipatory effect", in their terminology), Recasens and colleagues compared the steady state pattern of the V1 with lingual contact patterns in each EPG frame between the V1 steady state and the consonant onset. For calculating the C-on-V2 effect ("C-to-V carryover effect", in their terminology), Recasens et al. compared the steady state pattern of the V2 with lingual contact patterns in each EPG frame between the V2 steady state and the consonant offset. The largest difference between the vowel steady state lingual contact pattern and the lingual contact pattern elsewhere in that vowel was considered to be the size of the C-on-V effect. The researchers noted that "this maximal size difference usually occurs near closure onset for the anticipatory C-to-V effects and near closure offset for the carryover C-to-V effects" (Recasens et al. 1997, p. 548).

The DAC model, in its latest version (cf. Recasens 2004, in relation to Recasens et al. 1997), allows for the DAC values to be dependent on suprasegmental factors. This development of the model is following the predictions by Bladon and Al-Bamerni (1976), who claimed that CR values vary depending on both segmental and prosodic context. But the phenomena studied with EPG within the DAC model and the phenomena studied in this work are different. In Recasens (2004), consonant clusters were analysed, and the research question was whether syllable initial versus syllable final position affects consonant production. In this work, the question was whether the tautosyllabic vowel from a V#CV sequence had more influence on the consonant than the other vowel.

The differences in methodological details between this work and the EPG studies within the DAC model have the following implications. In order to describe the coarticulatory patterns observed in the ultrasound data in this work, the terminological apparatus of the DAC model is not sufficient. The argument is presented below.

V-on-C and C-on-V coarticulation was measured in this work using the same criteria, and the V-on-C coarticulation turned out to be significantly greater than the C-on-V coarticulation. This fact should be reflected in the theoretical representation of articulatory constraints of consonants and vowels. In Recasens et al. (1997), vowels and consonants are represented on the same DAC scale. For example, /a/ and /t/ are both characterised by the DAC value “2”, because the tongue dorsum is not directly involved in their production. A significantly greater V-on-/t/ coarticulation than C-on-/a/ coarticulation cannot be represented using this scale. There was another finding in this study that cannot be represented using the existing DAC value scales (neither the one presented in Recasens et al., 1997, nor the one presented in Recasens, 2004). This work demonstrated the influence of syllable boundary in the VCV sequence on the consonant resistance to the two vowels of the VCV. A different degree of vocalic influence on the consonant in a VCV sequence depending on the syllable affiliation of the consonant does not have a theoretical representation in the existing scales within the DAC model.

The previous paragraph suggests that three different types of phenomena observed in this study should be taken into account in the theoretical representation of articulatory constraints: resistance of the intervocalic consonant to the influence of the neighbouring vowels, resistance of the vowels to the influence of the consonant, and differerent degrees of resistance of the consonant to the two vowels, depending on syllable boundary. The proposed way of representing all these phenomena involves the term “degree of CR”, referring to the observed differences in tongue contour shapes. It is suggested that the degree of CR of a sound varies inversely with the distance between the tongue contours for this sound in different contexts. For example, when we compare the tongue contours for /h/ in different vowel contexts (Figure 5-4), the tongue contours for /t/ in different vowel contexts (Figure 6-7), and the tongue contours for /a/ in different consonant contexts (Figures 6-9 and 6-10), these three sounds can be arranged in ascending order, according to the degree of CR.

Using the concept of the degree of CR, we can describe some results of this work as follows. The degree of CR of non-lingual consonants (i.e., bilabial stops and /h/), lingual consonants and vowels changes in ascending order. Non-lingual consonants have the lowest degree of CR, lingual consonants have a higher degree of CR, and vowels have a much higher degree of CR than both types of consonants. It is possible to account for the stronger coarticulation within a CV-complex than within a VC-complex by assigning to the intervocalic consonant a higher CR degree in relation to V1 than to V2.

The model should also incorporate cross-language variation. At present, it is not explicitly incorporated in the DAC model. However, its influence on CR of speech sounds was predicted by Bladon and Al-Bamerni (see Section 2.1.1). Cross-linguistic differences observed in this work, in tongue behaviour in VCVs with non-lingual consonants, should be represented in the model. Using the concept of degree of CR should make it possible to account for all these differences.

In the next section, formulae are presented for quantifying degrees of CR for some of the different cases described above. This quantification, based on the actual numbers in the results, will allow for unifying the description of the vocalic influence on consonants, consonantal influence on vowels, and syllable boundary influence on VCV coarticulation.

7.3. Suggestions for quantifying resistance to lingual coarticulation, based on the ultrasound data from this work

In this section, some suggestions for quantifying CR of speech sounds are presented, with examples. The formulae introduced below are based on the type of data obtained in this work.

From the CR perspective, in this work there were two principal ways of displaying and measuring the tongue contour of a sound. One of them shows how much the sound is resistant to coarticulation with neighbouring sounds. On the graph, ultrasound curves for the target sound are presented together with the ultrasound curves for the neighbouring sounds. Examples include Figures 5-5 – 5-7, where the tongue contours during the production of British English /h/ are displayed together with the tongue contours for the surrounding vowels; Figures 6-3 – 6-5, where the tongue contours for V1, C and V2 are displayed on the same graph during the production of the British English /aka/, /ata/ and /iti/ sequences. This can be regarded as a representation of syntagmatic characteristics of the sound, i.e. how similar the sound is to the neighbouring sounds.

Another way of displaying and measuring the tongue contour of a sound shows how much the sound is coarticulated with neighbouring sounds. On the graph, ultrasound curves of the same sound in different contexts are presented. Examples include Figure 5-4, where the tongue contours during the production of British English /h/ are displayed in three different vowel environments; Figure 6-7, where the tongue contours during the production of the British English /t/ are displayed in two different vowel environments; Figures 6-9 and 6-10, where the tongue contours during the production of the British English /a/ are displayed in two different consonant environments. This can be regarded as a representation of paradigmatic characteristics of the sound, i.e. how similar the sound is in different contexts.

These two types of representation give us information on the two sides of the same phenomenon – interrelation of a speech sound with neighbouring sounds. One side is coarticulation, i.e., how dependent the sound is on others; the other side is resistance to coarticulation, i.e., how much the sound can retain its own identity. An attempt is made below to combine these two aspects in a unified measure.

7.3.1. Quantifying coarticulation resistance of a consonant

An example of the measure is based on the British English consonant /t/ produced by subject S2. The first way of representing coarticulatory properties concerns how much the consonant /t/ retains its own identity. The average curve of the consonant /t/ from the /iti/ sequences is displayed in Figure 7-1a, together with two /i/ curves from the /iti/ sequences; the average curve of the consonant /t/ from the /ata/ sequences is displayed in

Figure 7-1b, together with two /a/ curves. The distance between the C curve and the vowel curves is proportionate to the degree of resistance of /t/ to coarticulation.

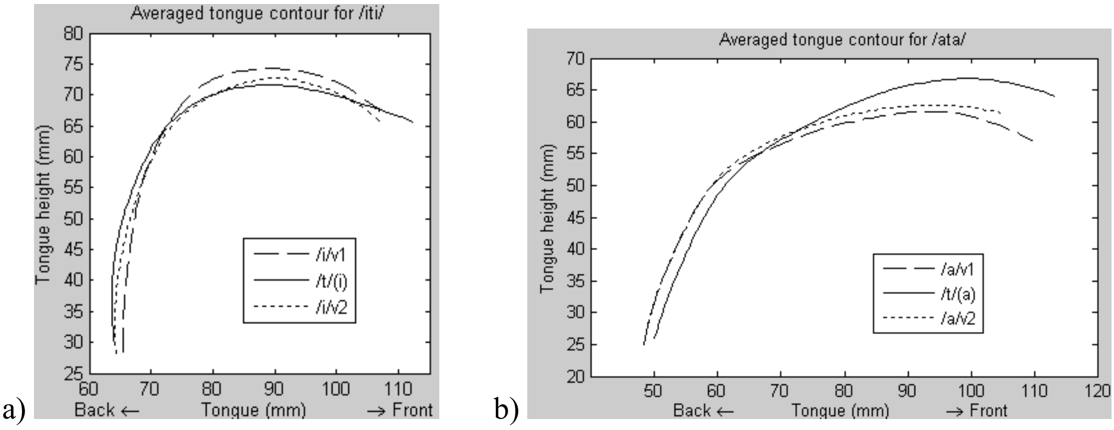


Figure 7-1. Average tongue contours in subject S2: a) in /iti/ sequences; b) in /ata/ sequences. Dashed line – V1 curve; solid line – C curve; dotted line – V2 curve.

The second way of representing this sound's properties is displayed in Figure 7-2, where two sets of tongue curves of 15 repetitions of the consonant /t/ are plotted: one set in the context of the vowel /i/, and the other set in the context of the vowel /a/. The distance between these two sets of curves is a measure of how much this sound is coarticulated. So this distance is in inverse proportion to the degree of resistance of /t/ to coarticulation.

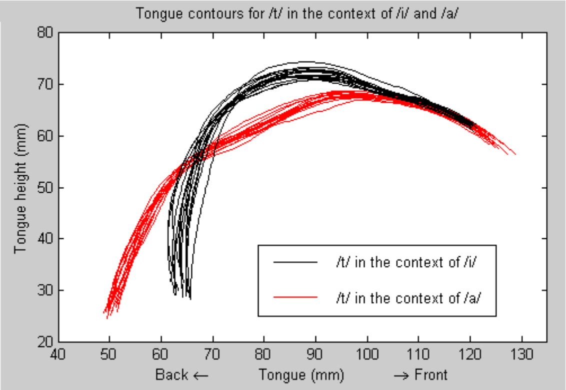


Figure 7-2. Tongue contours for fifteen repetitions of /t/ in subject S2, in two vowel environments: black lines – in the context of /i/; red lines – in the context of /a/.

The calculations are based on individual curves, as represented in Figure 7-2. Average curves are displayed in Figure 7-1, because it is difficult to see three sets of curves on the same graph when they are rather close together.

Given the representation of the coarticulatory properties of the consonant /t/ described in this section, the numerical value that we will obtain in the calculations will represent CR of the consonant in relation to these two vowels, /i/ and /a/, in VCV sequences. Below, the formula is presented, which describes CR of the intervocalic consonant /t/ in a VtV sequence, in relation to two different vowel contexts. The Coarticulation Resistance Coefficient (CRC) is a number representing CR of a speech sound in relation to the surrounding sounds. In this formula, the CRC of the consonant /t/ from a VtV sequence represents the CR of the consonant in relation to the surrounding vowels, /i/ and /a/. Here, the CRC is calculated in relation to both vowels from a VCV sequence (calculations can also be performed in relation to V1 and V2 separately; for details, see Section 7.3.3).

The data used for the calculations are 15 tokens of /iti/ and 15 tokens of /ata/ produced by one speaker. The distances from the consonant to its surrounding vowels (V1-t and t-V2) are proportionate to the degree of CR of the consonant, i.e., the degree to which /t/ retains its identity in a VCV sequence (see Figure 7-1). The V1-t and the t-V2 distances are computed within token, separately for each of the 15 /iti/ tokens and for each of the 15 /ata/ tokens. These distances are presented in Table 7-1. The average of these 60 distances is abbreviated as "C-V", and this value goes into the numerator of the formula below. The average of the distances in Table 7-1 is 2.8448.

The distance between the curves of the same consonant in two different vowel environments is in inverse proportion to the degree of CR of the consonant (see Figure 7-2). This distance is computed across all tokens, by comparing the 15 t_i curves with the 15 t_a curves (see Section 3.4.7.1 for the description of the procedure for computing this value). This distance is abbreviated in the formula as "C_i-C_a". This value goes into the denominator of the formula. For t_i and t_a curves, the distance is 4.6975 (see the matrix with 225 across-curve distances between t_i and t_a curves in Appendix VI-1).

iti		ata	
V1-t	**t-V2**	**V1-t**	**t-V2**
2.3626	1.4370	3.5233	2.6421
1.7108	2.2613	3.9155	2.9452
2.4763	1.8601	3.7131	2.6282
2.2037	1.6586	4.4877	3.7750
2.5960	2.0849	3.0781	2.4771
2.0076	1.5334	4.5045	3.8607
2.4022	2.3973	5.0397	3.6096
2.2820	2.5495	4.2967	3.2816
3.0398	2.4583	3.7717	2.9290
2.5199	1.7424	4.0388	2.7044
2.5240	1.8640	3.4991	3.2719
1.7305	1.6164	3.8639	3.3987
2.7253	1.8848	4.0723	3.4372
2.3234	1.4885	3.4545	2.8864
2.2964	1.4627	4.2072	3.8733

Table 7-1. V1-t and t-V2 distances, in millimetres, for obtaining the "C-V" value, for calculating the $CRC_{t(i,a)}$ for subject S2. Each row represents one /iti/ token and one /ata/ token.

The resulting value is multiplied by ten, for convenience. The formula is as follows:

$$CRC_{C(i,a)} \qquad \frac{(C\text{-}V) * 10}{C_i\text{-}C_a}$$

The calculations are illustrated below.

$$CRC_{t(i,a)} = \frac{2.8448 * 10}{4.6975} = 6.0560$$

The resulting CRC value is a ratio, it is not a number in millimetres. For convenience, it can be rounded to the nearest whole number. Thus, the $CRC_{t(i\text{-}a)}$ equals 6.

This value represents the degree of CR of the intervocalic /t/ to vocalic coarticulation, based on two particular vowel contexts.

In Table 7-2, the V1-h and h-V2 distances are given for the British English consonant /h/ from the contexts of /i/ and /a/, produced by the same subject. The average of the distances in Table 7-2 is 0.6091. The "C_i-C_a" value for /h/ equals 8.3175 (see the matrix with 225 across-curve distances between h_i and h_a curves in Appendix VI-2).

ihi		**aha**	
V1-h	**h-V2**	**V1-h**	**h-V2**
0.6837	0.5359	0.5209	0.3750
0.7698	0.6150	0.5733	0.3782
0.8580	1.1271	0.6394	0.4006
0.8027	0.7052	0.7954	0.4824
0.8454	0.6818	0.8030	0.4304
0.7359	0.6262	0.6705	0.4595
0.6142	0.5587	0.6007	0.4196
0.9304	0.6362	0.7704	0.4734
0.5738	0.3830	0.6642	0.4542
0.6825	0.5730	0.7144	0.4327
0.4802	0.3806	0.8366	0.3350
0.6451	0.6701	0.5805	0.5276
0.4516	0.5821	0.5299	0.3655
0.6208	0.6249	0.5269	0.4531
0.8074	0.9538	0.6024	0.5732

Table 7-2. V1-h and h-V2 distances, in millimetres, for obtaining the "C-V" value, for calculating the $CRC_{h(i,a)}$ for subject S2. Each row represents one /ihi/ token and one /aha/ token.

The calculations for /h/ produce the CRC number 0.7323, which rounds to 1. This tells us that, according to this method of calculation, the consonant /t/ is six times more resistant to vocalic coarticulation than the consonant /h/.

There is a possibility of combining this quantification with observations about specification of tongue parts, made throughout this work. In all the three experiments, three functional parts of the tongue have been mentioned: root, dorsum and blade. This division is somewhat arbitrary. Other functional divisions of the tongue have been proposed in the literature (see Section 6.4.4). Here, the division of the tongue into three parts will be used, as it turned out to be useful in describing and explaining the results of the experiments.

For British English /h/, only the root was shown to be specified (Experiment 2, Chapter 5), and for British English /t/, only the blade (Experiment 3, Chapter 6). In Table 7-3, these specifications are schematically represented.

	Root	Dorsum	Blade
h	+	–	–
t	–	–	+

Table 7-3. Specification of tongue parts in two British English consonants, /h/ and /t/, according to the results of this work. The sign "+" means that the tongue part is specified for the production of this sound; the sign "–" means that the tongue part is not specified for the production of this sound.

Combining the CRC values of the consonants /h/ and /t/ with the specifications from Table 7-3, it is possible to make some suggestions about the coarticulatory mechanisms in VCV sequences. For example, the CRC value of the consonant may be "residing" in the specified part of the tongue. This would allow us to predict that during the articulation of the consonant, the specified part of the tongue would actively move towards its position, and the remaining part(s) of the tongue would follow. The specified part of the tongue would retain its position with the strength defined by the CRC value of the consonant, and the other part(s) of the tongue would adapt to the neighbouring sounds more than the specified part.

Of course, in order to be able to predict the shapes of the tongue during the interaction of speech sounds, we should know not only which part of the tongue is specified for the production of these sounds, but also what the specification is. For the consonant /h/, according to our results, the root can be specified as retracted. For the consonant /t/, the blade can be specified as raised and touching the alveolar ridge. Here, it seems useful to mention the concept of markedness described in Section 2.3.1. If a part of the tongue is specified for a certain posture, this posture can be regarded as one of the members of the opposition, consisting of two or more different postures. For example, retracted tongue root for /h/ implies that there can be other postures for the tongue root. Certainly, these postures are best described in actual numbers (e.g., in millimetres) from the data. But for a schematic representation of possible specifications of the root position, we can accept that there are, e.g., three different postures of the root: retracted, advanced, and one in-between, that can be called neutral, and should probably correspond to the tongue root posture in the inter-uterance speech rest position (see Section 2.2.2). The same principles would apply to the dorsum and the blade. According to our data, the dorsum and the blade would probably each have to be part of two oppositions: one with the values “low”, “neutral” and “high” (vertical dimension), and the other one with the values “retracted”, “neutral” and “advanced” (horizontal dimension). The unmarked position for all the tongue parts would be neutral, and all the other postures would be marked.

The data presented in this work has been discussed in terms of three regions of the tongue, along the midsagittal line. It would be an interesting question for future research to identify other functionally important tongue regions. For example, there may be further subdivisions along the midsagittal line (e.g., Stone 1991; but see Nguyen et al. 1996). Evidence for the independent control of lateral parts of the tongue could be expected, based on Gibbon (1999). The information about lateral parts of the tongue may be a useful addition to midsagittal data (for example, in describing the vowel /i/ studied in Experiment 3 in this work). For this purpose, EPG would be a useful technique, as it displays patterns of contact of the tongue sides with the hard palate. Also, coronal ultrasound scanning could be used for the purpose of defining CRC specifications of the lateral parts of the tongue.

7.3.2. Quantifying coarticulation resistance of a vowel

An example of the quantification of CR of a vowel is based on the British English vowel /a/ produced by subject S2. As was the case for calculating the CRC of consonants, calculations for vowels are also performed in relation to context. In this work, /aka/ and

/ata/ sequences were studied, and these data are used for exemplifying the calculations. The CRC is calculated separately for V1 and V2.

One way of representing coarticulatory properties of the vowel shows how much the vowel retains its own identity. The average curves of V1 and V2 from the /aka/ sequences are displayed in Figure 7-3a, together with the average /k/ curve from the /aka/ sequences; the average curves of V1 and V2 from the /ata/ sequences are displayed in Figure 7-3b, together with the average /t/ curve from the /ata/ sequences. The distances between the vowel curves and the C curve are proportionate to the degree of resistance of /a/ to coarticulation.

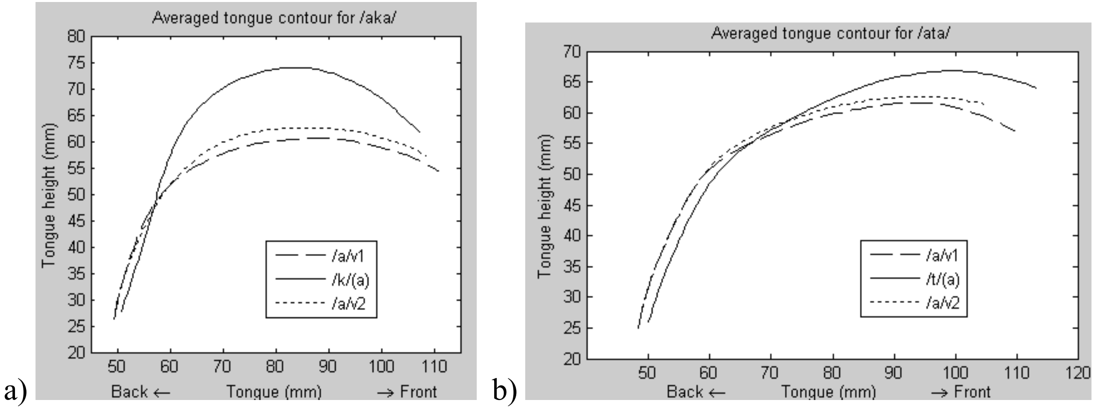


Figure 7-3. Average tongue contours in subject S2: a) in /aka/ sequences; b) in /ata/ sequences. Dashed line – V1 curve; solid line – C curve; dotted line – V2 curve.

The second way of representing this sound's properties is displayed in Figure 7-4, for the V1 of a VCV. Two sets of tongue curves of 15 repetitions of the vowel /a/ are plotted: one set in the context of the consonant /k/, and the other set in the context of the consonant /t/. The distance between these two sets of curves stands for how much this sound is coarticulated. So this distance is in inverse proportion to the degree of resistance of /a/ to coarticulation.

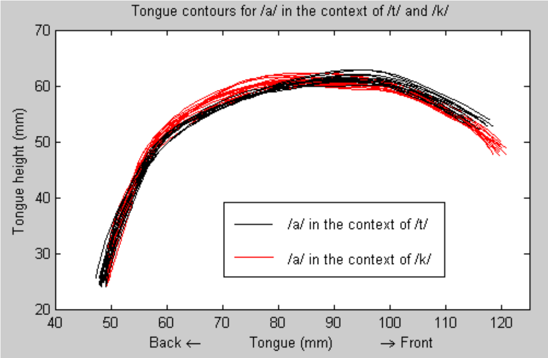


Figure 7-4. Tongue contours for fifteen repetitions of the first vowel /a/ in /aCa/ sequences, in subject S2, in two consonant environments: black lines – in the context of /t/; red lines – in the context of /k/.

Given the representation of the coarticulatory properties of the vowel /a/ described in this section, the numerical value that we will obtain in the calculations will represent CR of the vowel in relation to these two consonants, /k/ and /t/, in /aCa/ sequences. Below, the formula is presented, which describes the CR of the vowel /a/ in an /aCa/ sequence, in relation to two different consonant contexts. In this formula, the CRC of the vowel /a/ from an /aCa/ sequence represents the CR of the vowel in relation to the intervocalic consonant (/k/ or /t/). The formula is exemplified for the V1 of a VCV sequence. The same calculations can be performed for V2.

The data used for the calculations are 15 tokens of /aka/ and 15 tokens of /ata/ produced by one speaker. The distance from the vowel to the consonant is proportionate to the degree of CR of the vowel (see Figure 7-3). The a1-C distance is computed within token, separately for each of the 15 /aka/ tokens and for each of the 15 /ata/ tokens. The values for the first vowel of the VCV are presented in Table 7-4. The average of these 30 values represents the degree of resistance of the vowel to the consonantal influence. This average distance between the vowel and the neighbouring consonant is abbreviated as “V-C”, and this value goes into the numerator of the formula below. The average of the distances in Table 7-4 is 5.0658.

a1-C	
aka	**ata**
6.3515	3.5233
5.8371	3.9155
6.4772	3.7131
6.0080	4.4877
5.5893	3.0781
5.3246	4.5045
5.6907	5.0397
6.7640	4.2967
6.7686	3.7717
5.9290	4.0388
6.1618	3.4991
5.5869	3.8639
6.5560	4.0723
6.5801	3.4545
6.8830	4.2072

Table 7-4. Distances, in millimetres, for obtaining the "V-C" value, for calculating the $\mathrm{CRC}_{a1(k,t)}$ for subject S2. Each row represents one /aka/ token and one /ata/ token.

The distance between the curves of the same vowel in two different consonant environments is in inverse proportion to the degree of CR of the vowel (see Figure 7-4). This distance is computed across all tokens, by comparing the 15 ak curves with the 15 at curves (see Section 3.4.7.1 for the description of the procedure for computing this value). This distance is abbreviated in the formula as "V_k-V_t". This value goes into the denominator of the formula. For a1k and a1t curves, the distance is 1.1727 (see the matrix with 225 across-curve distances between a1k and a1t curves in Appendix VI-3).

The resulting value is multiplied by ten, for convenience. The formula is as follows:

$$\mathrm{CRC}_{V(k,t)} \quad \frac{(V\text{-}C) * 10}{V_k\text{-}V_t}$$

The calculations are illustrated below.

$$\mathrm{CRC}_{a1(k\text{-}t)} = \frac{5.0658 * 10}{1.1727} = 43.1977$$

The CRC value, rounded to the nearest whole number, is equal to 43. This value represents the degree of resistance of the first vowel from an /aCa/ sequence to the coarticulatory influence of an intervocalic consonant, based on two particular consonant contexts.

Table 7-5 contains the numbers required for the calculation of the CRC value for the V2 from the contexts of /k/ and /t/, produced by the same subject. The average of the distances in Table 7-5 is 4.0974. The "V_k-V_t" value for the V2 is equal to 1.3293 (see the matrix with 225 across-curve distances between $a2_k$ and $a2_t$ curves in Appendix VI-4). The $CRC_{a2(k-t)}$ is equal to 30.8237; rounded to the nearest whole number, the value is 31.

The CRC values for V1 and V2 are quite different. This is explained by the syllable boundary influence on the V1-C and C-V2 coarticulation: there is a stronger coarticulation of the consonant with the vowel belonging to the same syllable.

a2-C	
aka	**ata**
5.1419	2.6421
4.4905	2.9452
5.1761	2.6282
4.2408	3.7750
5.1955	2.4771
4.3430	3.8607
5.1062	3.6096
5.6815	3.2816
5.8277	2.9290
4.8388	2.7044
4.7461	3.2719
4.7477	3.3987
5.2520	3.4372
5.0111	2.8864
5.4016	3.8733

Table 7-5. Distances, in millimetres, for obtaining the "V-C" value, for calculating the $CRC_{a2(k,t)}$ for subject S2. Each row represents one /aka/ token and one /ata/ token.

Even though we observe a difference in CRC for V1 and V2, the CRC values for both vowels of a VCV sequence are much greater than the values we obtained for the consonants. This exemplifies the different mechanisms of vowel and consonant production referred to in Section 6.4.1. It was suggested in Section 6.4.1 that specification of tongue parts is rather different in consonants and vowels, in that the degree of CR of the unspecified part of the tongue is greater in vowels than in consonants. It was shown in Section 7.3.1 that only one part of the tongue is specified for the production of the consonants we studied. Specifications of tongue parts for the vowels studied in this work are discussed below. Our experimental data and the data from the literature (e.g., Browman & Goldstein 1990; Recasens et al. 1997) suggest that the tongue root is definitely specified for a certain position during the vowel production: namely, for an advanced position in /i/, and for a retracted position in /a/. The tongue dorsum is certainly specified for a raised and advanced position in /i/. It has been claimed in the literature that in front vowels, the front third of the tongue ("predorsum") is actively raised, while it stays low and inactive for the production of back vowels (Recasens 2002a). There are no indications in our results that the dorsum is specified for the production of the vowel /a/. Small differences were observed in the dorsum position in the two consonant contexts (/k/ and /t/); another fact is that in /aka/ sequences, there was quite a big displacement of the tongue dorsum upwards from V1 to /k/ and downwards from /k/ to V2. But these facts may be a consequence of the root pulling the dorsum downwards for the vowel production. The tongue blade appears to be not constrained for the vowel production, at least in the data presented in this work there is no evidence that the blade is specified for a certain position for the vowel production. So our data suggest that the vowel /a/ is not specified for a dorsum and a blade position, rather, the retraction of the root pulls the rest of the tongue downwards and backwards.

Table 7-6 presents schematic representations of the specification of tongue parts in the vowels mentioned in this section.

	Root	Dorsum	Blade
i	+	+	–
a	+	–	–

Table 7-6. Specification of tongue parts in two British English vowels, /i/ and /a/, according to the results of this work. The sign "+" means that the tongue part is specified for the production of this sound; the sign "–" means that the tongue part is not specified for the production of this sound.

Greater CRC values for vowels than for consonants, combined with a specification of more parts of the tongue in some vowels than in consonants, reflect the observed coarticulatory patterns in VCV sequences. Greater CRC values for vowels than for consonants suggest that not only the specified part(s) of the tongue, but also the rest of the tongue contour in vowels is much more resistant than in consonants. An example of the interaction of a vowel with a consonant is /ihi/ sequences. The vowel /i/ and the consonant /h/ have contrasting tongue root position specifications: the vowel is specified for an advanced tongue root, while the consonant is specified for a retracted tongue root. The CRC of the vowel is much greater than the CRC of the consonant. The coarticulatory pattern reported in Experiment 2 (Chapter 5) is characterised by a relatively advanced tongue root for both the C curve and the vowel curves, and generally the C curve has a very similar shape to the vowel curves. The tongue root is slightly retracted in the C curve, in relation to the vowel curves. This slight tongue root retraction in /h/ is a demonstration of the consonant's CR, which is much smaller than that of the vowel.

7.3.3. Quantifying coarticulation resistance of a consonant in relation to syllable boundary

One of the findings in this work was that CR of the sounds forming a VCV sequence depends on the place of the syllable boundary. An example of quantification of the CR of the intervocalic consonant in relation to the syllable boundary in a VCV sequence is presented below.

The formula for calculating the CRC of a consonant in relation to syllable boundary is the same as the formula described in Section 7.3.1, which produces the CRC value of a consonant in relation to both vowels of a VCV. For representing syllable boundary influence on the consonant's CR, the CRC for the intervocalic consonant is calculated separately in relation to V1 and to V2. The example of the calculations is given for the British English consonant /t/, in relation to V1, based on two vowel contexts, /i/ and /a/.

The data used for the calculations are the same as those used in Section 7.3.1, i.e., 15 tokens of /iti/ and 15 tokens of /ata/ produced by one speaker. The "C-V" value is obtained as follows. The V1-t distance is computed within token, separately for each of the 15 /iti/ tokens and for each of the 15 /ata/ tokens. The values are presented in Table 7-7. The average of these 30 distances is proportionate to the degree of resistance of /t/ to the influence of V1 in a VCV sequence. The average of the distances in Table 7-7 is 3.1556.

V1-t	
iti	**ata**
2.3626	3.5233
1.7108	3.9155
2.4763	3.7131
2.2037	4.4877
2.5960	3.0781
2.0076	4.5045
2.4022	5.0397
2.2820	4.2967
3.0398	3.7717
2.5199	4.0388
2.5240	3.4991
1.7305	3.8639
2.7253	4.0723
2.3234	3.4545
2.2964	4.2072

Table 7-7. Distances, in millimetres, for obtaining the "C-V" value, for calculating the $CRC_{t(i1,a1)}$ for subject S2. Each row represents one /iti/ token and one /ata/ token.

The "C_i-C_a" value is the same as the one used in Section 7.3.1, it equals 4.6975 (see the matrix with 225 across-curve distances between t_i and t_a curves in Appendix VI-1). The calculations are illustrated below.

$$\mathrm{CRC}_{t(i1,a1)} = \frac{3.1556 * 10}{4.6975} = 6.7176$$

Rounded to the nearest whole number, the $\mathrm{CRC}_{t(i1,a1)}$ is equal to 7. This value represents the degree of resistance of the intervocalic /t/ to the influence of the first vowel of a VCV sequence, based on two particular vocalic contexts.

Table 7-8 contains the t-V2 distances, required for the calculations of the CRC value for /t/ in relation to V2. The average of the distances in Table 7-8 is 2.5340. The "C_i-C_a" value is 4.6975 (see Appendix VI-1). The $\mathrm{CRC}_{t(i2,a2)}$ is equal to 5.3944; rounded to the nearest whole number, the value is 5.

We can see that resistance of /t/ to the second vowel of a VCV sequence is smaller than its resistance to the first vowel. When the CRC values are rounded to nearest whole numbers, we obtain the values of 5 and 7 for the resistance to V2 and V1, respectively. This quantification illustrates the coarticulatory patterns observed in Experiment 3.

t-V2	
iti	**ata**
1.4370	2.6421
2.2613	2.9452
1.8601	2.6282
1.6586	3.7750
2.0849	2.4771
1.5334	3.8607
2.3973	3.6096
2.5495	3.2816
2.4583	2.9290
1.7424	2.7044
1.8640	3.2719
1.6164	3.3987
1.8848	3.4372
1.4885	2.8864
1.4627	3.8733

Table 7-8. Distances, in millimetres, for obtaining the "C-V" value, for calculating the $\mathrm{CRC}_{t(i2,a2)}$ for subject S2. Each row represents one /iti/ token and one /ata/ token.

7.3.4. Representation of coarticulation resistance of speech sounds in the enlarged DAC model

It was mentioned in Recasens et al. (1997) that the ternary classification of degrees of articulatory constraint is preliminary, and could be improved when the articulatory constraints for consonants and vowels are formulated more accurately. In later publications, Recasens proposes some changes in the classification, with not only the tongue dorsum, but also the “tongue front” having the possibility to be constrained (see Section 2.1.2 for details). The data reported in the present study suggest that the tongue root should also be included in the classification of degrees of articulatory constraint.

In Table 7-9, there is a schematic representation of lingual CR characteristics of the consonants /h/ and /t/ and the vowel /a/ in VCV sequences, based on the data from subject S2. These segments are chosen for illustrative purposes, as they were used in this chapter for calculating the CRC values. The representation in Table 7-9 is based on the ultrasound data from this work, but it also includes the classification from Recasens et al. (1997). The assignment of the DAC value is based on the involvement of the tongue dorsum in the production of the sound (for more details on the DAC classification, see Recasens et al. 1997). CRC values are calculated based on comparing two different types of VCV sequences: /aha/ and /ihi/ for calculating the CRC of the consonant /h/; /ata/ and /iti/ for calculating the CRC of the consonant /t/; /aka/ and /ata/ for calculating the CRC of the vowel /a/. For explanations of the details concerning the calculation of CRC values, see Sections 7.3.1 – 7.3.3.

In the table, the first column contains the speech sounds that are being described. The second column has the DAC value, which is assigned to the sound based on the involvement of the tongue dorsum in the production of the sound. The third, the fourth and the fifth column represent specification of the tongue parts for producing the sound. The “–” sign means that a tongue part is not specified for a particular position for producing the sound. If a tongue part is specified for a certain position for producing the sound, schematic specifications are given in the cell, according to the criteria suggested in Section 7.3.1. The sixth column in the rows for /h/ and /t/ contain the CRC values of these consonants, as calculated in Section 7.3.1. The sixth column in the row for /a/

contains an average value over the CRC values calculated for V1 and V2 in Section 7.3.2. The last two columns in the rows for /h/ and /t/ represent the CR of the consonant to the V1 and the V2, respectively (see Section 7.3.3). The last two columns for the vowel /a/ represent the CR to the intervocalic consonant of the V1 and of the V2, respectively (see Section 7.3.2).

	DAC value	Root	Dorsum	Blade	CRC	CRC_{V1}	CRC_{V2}
/h/	1	retracted	–	–	1	0.8	0.6
/t/	2	–	–	advanced & raised	6	6.7	5.4
/a/	2	retracted	–	–	37	43	31

Table 7-9. Schematic representation of lingual CR characteristics of the consonants /h/ and /t/ and the vowel /a/ in VCV sequences, based on the ultrasound data from subject S2.

This representation allows for distinguishing between the consonant /t/ and the vowel /a/, which are characterised by the same DAC value in the classification presented in Recasens et al. (1997). Also, there is information about involvement of different tongue parts in sound production. Besides, the influence of the syllable boundary on the CR of the sounds is represented.

This way of representing articulatory constraints, based on the ultrasound data from this research, does not contradict the ternary classification of degrees of articulatory constraint presented in Recasens et al. (1997), nor the one presented in Recasens (2004). Rather, it can be seen as a more detailed representation of articulatory constraints.

7.4. Final observations on the CR approach and the DAC model in relation to this study

The CR approach to speech production and the DAC model have been useful and convenient in this research. However, this work demonstrated that the infrastructure of the DAC model could not accommodate all the ultrasound results. Some changes have been introduced in the DAC model (Section 7.3). The modified version of the model allows for representing articulatory constraints of speech sounds in more detail.

One interesting issue arising from the comparative analysis of this work and previous studies made within the DAC framework is that any measurement of coarticulation is relative, because isolated speech sounds do not exist. The results of all measurements depend on various factors, e.g., which speech segment is studied, what the context is, what the speech rate is, whether the speaker has any speech disorders, etc. For example, Recasens and his co-authors measured the C-on-V coarticulation by using the steady state of the vowel as a reference. However, the steady state of the vowel could have also been influenced by the consonant. The results of this work demonstrated that the tongue shape at the steady state of the vowel /a/ was significantly different in two different consonant contexts, /k/ and /t/. The results of Recasens et al. (1997) and the results of this work show different sides of C-on-V coarticulation. So these two sets of results can be regarded as complementary, enlarging the theoretical knowledge in different ways.

One more difference between earlier EPG studies within the DAC model and this work is that EPG and ultrasound produce complementary information on lingual articulation. Ultrasound provides the information about the location of the tongue in space, not merely about the contact to the hard structures of the palatal and alveolar passive articulators. For example, EPG does not allow us to observe C-on-V coarticulation in the vowel /a/ steady states. EPG is not a good method for comparing steady states of open vowels, because there is very little contact of the tongue with the palate during production of open vowels. In the present study, lingual contact patterns at the steady state of the vowel /a/ were similar in two consonant contexts (/k/ and /t/).

However, ultrasound data produced evidence of consonant-related differences in the tongue shapes at the vowel steady state. Using ultrasound together with EPG appears to look effective. The high frame rate of EPG allows for obtaining information on fine detail in articulator timing. The ability of EPG to provide data on lateral bracing, taken with the ability of ultrasound to display the whole tongue curve, can give us more precise information on the tongue location than we can get from any one of these techniques alone.

The question of cross-speaker variability can be raised in relation to the enlarged version of the DAC model proposed in this chapter. Given that CRC values are based on speech data, and not on predictions or estimates, there is some individual variation to be expected in CR characteristics of speech sounds. This question has not been explicitly addressed in this work, and it remains a challenge for future research to define how variable and how reliable CRC values are. Preliminary calculations for the three subjects in Experiment 3 (Chapter 6) show that in all the subjects vowels have much greater CRC values than consonants, and the CRC of the intervocalic consonant is smaller in relation to the second vowel of the VCV than to the first vowel.

It would be interesting to use CRC values for analysing other linguistic phenomena. For example, it was mentioned in Chapter 4 that stressed and unstressed vowels exhibit different degrees of coarticulation resistance. An experiment could be designed which would allow for quantifying stress influence on the CR of vowels. Another application of this technique would be to measure CR of the consonants considered ambisyllabic. CRC values of such consonants in relation to the two surrounding vowels would give us evidence about syllable affiliation of these consonants. Also, CRC values of speech sounds in relation to the position in a word or phrase could be calculated. A challenging application of the CR concept would be to design an approach that would combine ultrasound and EPG data for calculating CRC values. This could potentially be applied in cases where information on the lateral parts of the tongue is important, together with midsagittal data. One more application of this technique would be to quantify CR of parts of the tongue separately. This would give us more insights into mechanisms of motor control than we can get from the calculations based on the whole tongue contour.

The method for quantifying CR described in this chapter, as well as all the ideas for its application in analysing linguistic questions, could be used to study both non-disordered and abnormal speech production. A comparative study of CRC values in disordered and non-disordered speakers would help to identify, describe and predict differences between normal and clinical populations.

8. SUMMARY AND FINAL CONCLUSIONS

8.1. General summary

In this work ultrasound was used in order to study tongue behaviour during the production of VCV sequences. Ultrasound was shown to be a very efficient technique, providing informative and interesting results. The notion of Coarticulation Resistance (CR) was employed for interpreting experimental results. The Degree of Articulatory Constraint (DAC) model proved to be a convenient framework for a consistent description of the coarticulatory properties of speech sounds.

In Experiment 1, Russian symmetrical VCV sequences with bilabial stops and the vowels /i/, /u/, /a/, produced by three speakers, were examined. It was qualitatively shown that the tongue shape for bilabial consonants greatly depends on the tongue shapes of the surrounding vowels. This finding was interpreted by claiming that the CR of Russian bilabial stops in symmetrical VCV sequences was not absolute.

No tongue lowering (i.e., “trough” patterns) was found between the vowels in Russian VCV sequences with high vowels, contrary to the patterns observed in the literature for, e.g., English and Swedish. In /aCa/ sequences with bilabial consonants, tongue raising (“antitrough” patterns) between the two vowels was observed. These findings were explained by the fact that the phonological opposition of Russian consonants in palatalisation imposes a certain constraint on the tongue position for these consonants’ production. Namely, the tongue dorsum is raised, in relation to the neutral position, and located towards the front or the back of the mouth, for producing palatalised and non-palatalised consonants, respectively.

Qualitative comparison of the graphs featuring tongue contours for V1, C and V2, from the Russian VCVs with bilabial consonants, showed a tendency for the C curve to be closer to the V2 curve than to the V1 curve. This was interpreted as the influence of the syllable boundary on the VCV coarticulatory pattern, and on the degree of CR of the intervocalic consonant. Namely, the intervocalic consonant was suggested to have a lower degree of CR in relation to V2 than to V1.

Experiment 2 was aimed at studying lingual coarticulation in British English VhV sequences with the vowels /i/, /u/, /a/. As in Experiment 1, it was qualitatively shown that the vowels greatly influence the tongue shape for the consonant. A quantitative procedure was designed and applied, and it was demonstrated that the consonant's tongue contour differs significantly, depending on the vowel context. This result was taken to signify that CR of British English /h/ in symmetrical VCV sequences was not absolute.

Tongue root retraction and tongue dorsum lowering ("trough" patterns) were found between the two vowels in /ihi/ and /uhu/ sequences. A continuous tongue movement between the two vowels was observed in /aha/ sequences. It was suggested that troughs in VhVs with high vowels could be explained by intervocalic relaxation of suprahyoid muscles, for lowering the larynx, which was raised during the production of the two high vowels of a VhV. In /aha/ sequences, the larynx was not raised for the production of the vowels, hence no discontinuity in lingual coarticulation was observed. Within the DAC framework, it was suggested that the tongue root was specified for a retracted position for producing the consonant /h/.

A quantitative procedure was designed and applied to the comparison of V1-/h/ and /h/-V2 distances. The outcome of the analysis demonstrated that /h/-V2 distances were significantly smaller than V1-/h/ distances only in one of the two subjects analysed. These results were interpreted by suggesting that the British English /h/ does not necessarily have a lower CR to V2 than to V1 in the V#CV sequence.

In Experiment 3, coarticulatory patterns in /a#ka/, /a#ta/ and /i#ti/ sequences were studied. The tongue shape for the consonant was shown to vary, according to the surrounding vowels' tongue shapes. Quantitative analysis demonstrated a significant difference between the /t/ curves in the context of /a/ versus /i/. This was interpreted by suggesting that /t/ was not absolutely resistant in VCVs with these vowels. Qualitative comparison of VtV sequences from this experiment to VCVs from Experiments 1 and 2 showed that the surrounding vowels influenced the lingual shape for non-lingual consonants much more than for /t/. This was explained by differing requirements on tongue position for the consonant production: for producing /t/, the tongue is more constrained than for producing non-lingual consonants, because the blade needs to make an occlusion for /t/. These results allowed for describing /t/ as having a higher CR in VCV sequences than non-lingual consonants.

Coarticulatory influence of consonants on neighbouring vowels was also analysed. The sequences /ata/ and /aka/ were compared, and the tongue curves of the vowels were shown to differ significantly, depending on the consonant context. This was interpreted by claiming that CR of the vowel /a/ was not absolute in /aCa/ sequences with lingual consonants.

The influence of different vowels on the consonant /t/ was demonstrated to be significantly greater than the influence of different consonants on the vowel /a/. This finding was explained by the fact that for vowel production, the tongue is more constrained than for consonant production. Acoustic characteristics of consonants result from an occlusion or constriction made by a certain part of the tongue, while acoustic characteristics of vowels are more dependent on the whole shape of the tongue. These differing characteristics of consonants and vowels, demonstrated by the findings, were interpreted by suggesting that in VCV sequences, vowels have a higher degree of CR than consonants.

V1-C and C-V2 distances were compared for British English VCVs with lingual consonants, and significant differences (C-V2 distances being smaller) were found in most cases. This was interpreted as a manifestation of syllable and word boundary influence on the VCV coarticulatory pattern, and on the CR of the intervocalic consonant. The intervocalic consonant was suggested to have a lower CR in relation to V2 than to V1.

In general, the results of this work suggest that degrees of CR are highly relative. It has already been claimed in the literature within the DAC framework that gestural coordination of neighbouring segments depends on several segmental and prosodic factors (e.g., Recasens 2002a). Our results provide more evidence in support of this claim, and also enable us to suggest that degrees of CR can only be specified in relation to contextual segmental and prosodic interaction. CR was found to be dependent on the type of the segment (vowel or consonant), on the context, on the syllable boundary, and, in the case of consonants, on the place of articulation. Demands on the articulators coming from phonological oppositions in the language were also shown to affect the degree of CR.

Suggestions for quantifying CR of speech sounds were presented, with examples. The DAC model was critically evaluated and discussed in the light of the data presented in this work. Ways of enlarging the DAC model were suggested, in order to account for more aspects of speech production than it presently does. Specifically, functional division of the tongue into three parts was introduced: blade, dorsum and root. Lingual position specifications were described separately for each of these three tongue parts. The concept "degree of CR" was presented, allowing for a unified description of coarticulatory properties of consonants and vowels, and for representing syllable boundary influence on lingual coarticulation in VCV sequences. The Coarticulation Resistance Coefficient (CRC) was introduced for quantifying degrees of CR of speech sounds.

8.2. Conclusions on methodological aspects

At the beginning of the study, theoretical issues were overriden by technical complexities and challenges. The methods progressed technically, as the procedures for tongue contour measurement were elaborated in the course of the project. The solutions for the technical problems are described below.

The first technical challenge was the need for visualising the results. Methods for plotting tongue curves in xy coordinates were designed, as described in Section 3.4.5. Then, in order to visually display vowel-on-consonant coarticulatory influence, two strategies were used. One consists in making average three-curve graphs over 15 repetitions of the same VCV sequence, so that the V1 curve, the C curve and the V2 curve are displayed on the same graph (e.g., Figures 4-5 – 4-7). This type of graph, as was demonstrated in Chapters 4 and 5, produces a good impression of how close the consonant tongue curve is to the two vowels' tongue curves. The second type of graph features C curves from different vowel environments (e.g., Figure 5-7). This graph gives a clear visual representation of how different the consonant tongue curves are in three different vowel contexts (/i/, /u/ and /a/). Based on qualitative observations, or “eye-balling”, of these graphs, conclusions were made in Chapters 4 and 5 about the strong influence of the vowels on the intervocalic consonants, as evidenced by the ultrasound data.

One of the implications of the first experiment’s results was that a quantitative procedure was necessary, in order to numerically confirm the qualitative impressions of the differences in the consonant tongue shape depending on the vocalic environment. The procedure was then designed (see Section 3.4.7.2 for description), and proved to be successful in demonstrating that the consonant tongue curves for /h/ were significantly different in the three vowel environments (see Section 5.3.1). The same Matlab-based procedure was used for the analysis of non-lingual consonants and vowels, in Experiment 3, in order to define whether consonant curves differed across vowel environments, and whether vowel curves differed across consonant environments.

For analysing syllable boundary influence on VCV coarticulation, first it was necessary to find methods to visualise syllable boundary influence on the tongue shapes for V1, C and V2, and then to quantify this influence. In the first experiment, qualitative observations were made, based on average three-curve graphs over 15 repetitions of the same VCV. Then, as one of the outcomes of Experiment 1, there arose the need for a reliable quantitative procedure that would allow for numerically showing whether the syllable boundary effect was present in V#CV sequences. The procedure was designed (see Section 3.4.7.1 for details), and applied for data analysis in Experiments 2 and 3.

In general, ultrasound proved to be a useful technique for measuring lingual coarticulatory properties of speech sounds. Certain technical challenges of this method of studying articulation have to be taken into account when designing new experiments. For example, the quality of the ultrasound image highly depends on the vocal tract morphology of individual subjects, so there should be pre-test screenings, in order to ensure that the quality is acceptable. Also, different speech sounds image with different quality, depending on how far the tongue surface is from the transducer, so before running the experiment, it should be tested whether the target data can be satisfactorily imaged.

The methodology of combined ultrasound, EPG and acoustics was shown to be useful, providing supporting complementary evidence from EPG and ultrasound. On the basis of this study, a conclusion can be drawn that this multi-channel technique deserves further refinements, and designing combined quantitative analysis methods will be particularly useful, in order to implement this technique in further speech research.

8.3. Directions for future research

This work has demonstrated that a combination of the theoretical CR approach and the methodology involving ultrasound can be effective in analysing coarticulatory effects occurring in speech production. Therefore, future application of this theoretical framework and/or of this technique for studying speech production will certainly be very useful. Below, a few possible directions for future work are mentioned.

In this study, the data were VCV sequences, some of them nonsense VCVs, and others occurring in meaningful words and sentences. As the results show some interesting coarticulatory patterns in these sequences, it seems feasible to conduct a similar type of analysis on continuous speech data, and not lists of sentences. The data could be designed to incorporate certain sound sequences. If spontaneous speech is studied, the sound sequences of interest may be elicited from the subjects by using specific tasks, e.g., describing a picture.

An issue that arose in this work was some evidence towards specification of parts of the tongue (e.g., root, dorsum or blade) for producing certain sounds, as compared with the notion of the whole togue being specified for a particular position. It would be very interesting to use ultrasound for obtaining more data and numerically showing these specifications. Then it would be possible to describe lingual articulatory constraints of various speech sounds (or whole phoneme systems) using the extended version of the DAC model proposed in this work.

The theoretical approach adopted in this work, using the concept of CR for describing the observed coarticulatory patterns, proved to be plausible and helpful. The same approach could be used in future studies. It would be interesting to further develop the DAC model, in order to include in its numerical descriptions more factors than were included in the quantification of CR in this research: for example, word and phrasal position, word stress, sentence stress, cross-language and cross-speaker differences. Finally, it may be interesting to situate the DAC model within a larger theory. It may be based on the CR approach to speech production, but the crucial point would be that it would involve not only lingual position characteristics of speech sounds, but also information from other articulators, as well as information on acoustic and perceptual specification of speech sounds.

Quantifying and comparing degrees of CR across languages could be a very interesting addition to existing studies describing phonetic characteristics of the world's languages. Ultrasound could be applied to studying tongue shapes required for producing various segments in the languages of the world that have not been studied on a large scale.

An important outcome of this study was that it demonstrated the possibility of using ultrasound together with EPG, and obtaining reliably comparable results from the two techniques. In future, ultrasound and EPG could be used for analysing tongue contour data together with tongue-palate contact data. Another interesting future direction could be using ultrasound together with EMA, in order to be able to track individual flesh points of the tongue, and not only the general displacement of the tongue contour. In order to use these techniques together, a lot of preparatory calibration and synchronisation work is required, but then we could obtain potentially very interesting results.

A strength of ultrasound is that the research conducted with this technique can easily be applied in clinical work and for teaching purposes. As the procedure of ultrasound scanning is non-invasive and not very demanding of the subjects, research may be carried out with children, both normally developing and with a speech pathology, as well as with adults having speech disorders. Obtaining ultrasound articulatory data from normally developing children would be very useful in order to better understand the development of motor control. Normative data would make it easier to identify developmental articulatory disorders, and will be helpful for introducing ultrasound into clinical practice.

Ultrasound can be used in speech and language therapy and in the teaching of phonetics. It allows patients or students to instantly see both the correct position of the tongue during producing a sound (from a therapist or a teacher) and the feedback from their own attempts to produce it. Thus they get some visual stimulation and motivation for correcting themselves. It would be very interesting to apply ultrasound in future research for designing clinical treatment programmes and practical tools for teaching phonetics and linguistics.

REFERENCES

Al-Halees, Y. (2003). A xeroradiographic, ultrasonic, laryngoscopic and acoustic study of pharygeals and emphatics in Arabic. In *Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS)*, Barcelona, Spain. 1747-1750.

Amerman, J. D., Daniloff, R. G. & Moll, K. (1970). Lip and jaw coarticulation for the phoneme /æ/. *Journal of Speech and Hearing Research*, **13**, 147-61.

Archangeli, D. B. (1992). Markedness in phonology. In W. Bright (Ed.), *International Encyclopaedia of Linguistics*. New York, Oxford: Oxford University Press. Vol. 2. 391-393.

Bell-Berti, F. & Harris, K. S. (1974). More on the motor organization of speech gestures. *Haskins Laboratories Status Report on Speech Research*, **SR-37/38**, 73-77.

Bell-Berti, F. & Harris, K. S. (1979). Anticipatory coarticulation: some implications from a study of lip rounding. *Journal of the Acoustical Society of America*, **65**, 1268-1270.

Bell-Berti, F. & Harris, K. S. (1981). A temporal model of speech production. *Phonetica*, **38**, 9-20.

Bell-Berti, F. & Harris, K. S. (1982). Temporal patterns of coarticulation: lip rounding. *Journal of the Acoustical Society of America*, **71**, 449-454.

Benguerel, A.-P. & Cowan, H. A. (1974). Coarticulation of upper lip protrusion in French. *Phonetica*, **30**, 41-55.

Benus, S. (2005). *Dynamics and Transparency in Vowel Harmony.* PhD dissertation, New York University.

Bernhardt, B., Gick, B., Bacsfalvi, P. & Ashdown, J. (2003). Speech habilitation of hard of hearing adolescents using electropalatography and ultrasound as evaluated by trained listeners. *Clinical Linguistics and Phonetics*, **17**, 199-216.

Bernhardt, B., Bacsfalvi, P., Gick, B., Radanov, B**.** & Williams, R. (2005). Exploring electropalatography and ultrasound in speech habilitation. *Journal of Speech-Language Pathology and Audiology*, **29**, 169-182.

Bladon, R. A. W. & Al-Bamerni, A. (1976). Coarticulation resistance in English /l/. *Journal of Phonetics*, **4**, 137-150.

Boersma, P. (1998). Spreading in functional phonology. *Proceedings of the Institute of Phonetic Sciences, Amsterdam*, **20**, 1-20.

Bondarko, L. V. (1969). The syllable structure of speech and distinctive features of phonemes. *Phonetica*, **20**, 1-40.

Bondarko, L. V. (1998). *Fonetika sovremennogo russkogo jazyka*. (*Phonetics of Modern Russian.*) St. Petersburg: St. Petersburg University Press.

Bondarko, L. V. (2005). Phonetic and phonological aspects of the opposition of ‘soft’ and ‘hard’ consonants in the modern Russian language. *Speech Communication*, **47**, 7-14.

Bondarko, L. V., Verbitskaja, L. A. & Zinder, L. R. (1966). Akusticheskije harakteristiki bezudarnosti: na materiale russkogo jazyka. (*Acoustic characteristics of unstressed vowels: Russian data.*) In V. V. Ivanov (Ed.), Strukturnaja tipologija jazykov (*Structural Typology of Languages*). Moscow: USSR Academy of Sciences. 56-64.

Bondarko, L. V., Zinder, L. R. & Svetozarova, N. D. (1968). Razgranichenije slov v potoke rechi. (*Distinguishing words in continuous speech.*) *Voprosy jazykoznanija*, **2**, 68-81.

Bothorel, A., Simon, P., Wioland, F., Zerling, J.-P. (1986). *Cinéradiographie: des voyelles et consonnes du français. Traveaux de l'institut de phonétique de Strasbourg*. Strasbourg: Publications de l'institut de phonétique de Strasbourg.

Boyce, S. (1990). Coarticulatory organization for lip rounding in Turkish and English. *Journal of the Acoustical Society of America*, **88**, 2584-2595.

Bressman, T., Heng, C.-L. & Irish, J. C. (2005). Applications of 2D and 3D ultrasound imaging in speech-language pathology. *Journal of Speech-Language Pathology and Audiology*, **29**, 158-168.

Browman, C. P. & Goldstein, L. (1988). Some notes on syllable structure in articulatory phonology. *Phonetica*, **45**, 140-155.

Browman, C. P. & Goldstein, L. (1990). Tiers in articulatory phonology, with some implications for casual speech. In J. Kingston & M. E. Beckman (Eds), *Papers in Laboratory Phonology II: Between the Grammar and the Physics of Speech*. Cambridge: Cambridge University Press. 341-376.

Browman, C. P. & Goldstein, L. (1992a). Articulatory phonology: an overview. *Phonetica*, **49**, 155-180.

Browman, C. P. & Goldstein, L. (1992b). "Targetless" schwa: an articulatory analysis. In G. Docherty & D. R. Ladd (Eds), *Papers in Laboratory Phonology I: Gesture, Segment, Prosody*. Cambridge: Cambridge University Press. 26-56.

Brugman, J. (2005). Pharyngeal constrictions in the vowels of N|u. [Oral paper, Ultrafest III 2005 (Ultrasound meeting), Arizona Phonological Imaging Lab, University of Arizona, Tucson, AZ, USA.]

Byrd, D. M. (1995). C-Centers revisited. *Phonetica*, **52**, 263-282.

Byrd, D. & Saltzman, E. (2003). The elastic phrase: modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics*, **31**, 149-180.

Chomsky, N. & Halle, M. (1968). *The Sound Pattern of English*. New York, NY: Harper and Row.

Cohen, M. M., Beskow, J. & Massaro, D. W. (1998). Recent developments in facial animation: an inside view. [Paper at the International Conference on Audio-Visual Speech Processing (AVSP'98), Terrigal, Australia.]

Crystal, D. (1997). *A Dictionary of Linguistics and Phonetics*. 2nd edn. Oxford: Blackwell.

Davidson, L. (2005a). Addressing phonological questions with ultrasound. *Clinical Linguistics and Phonetics,* **19**, 619-633.

Davidson, L. (2005b). Syllabic position effects in articulation: are word-medial and word-final codas equivalent? [Oral paper, Ultrafest III 2005 (Ultrasound meeting), Arizona Phonological Imaging Lab, University of Arizona, Tucson, AZ, USA.]

References

Davidson, L. (2006a). Phonology, phonetics, or frequency: influences on the production of non-native sequences. *Journal of Phonetics*, **34**, 104-137.

Davidson, L. (2006b). Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *Journal of the Acoustical Society of America*, **120**, 407-415.

Davidson, L. & Stone, M. (2003). Epenthesis versus gestural mistiming in consonant cluster production: An ultrasound study. In G. Garding and M. Tsujimura (Eds), *WCCFL 22 Proceedings*. Somerville, MA: Cascadilla Press. 165-178.

Engstrand, O. (1981). Acoustic constraints or invariant input representation? An experimental study of selected articulatory movements and targets. *Reports from Uppsala University, Department of Linguistics (RUUL)*, **7**, 67-95.

Engstrand, O. (1988). Articulatory correlates of stress and speaking rate in Swedish VCV utterances. *Journal of the Acoustical Society of America*, **83**, 1863-1875.

Engstrand, O. (1989). Towards an electropalatographic specification of consonant articulation in Swedish. *Phonetic Experimental Research, Institute of Linguistics, University of Stockholm (PERILUS)*, **10**, 115-156.

Engstrand, O., McAllister, R. & Lindblom, B. (1996). Investigating the "trough": vowel dynamics and aerodynamics. *Journal of the Acoustical Society of America*, **100**, 2659-2660.

Ericsdotter, C., Lindblom, B. & Stark, J. (1999). Articulatory coordination in coronal stops: implications for theories of coarticulation. *Proceedings of the 14th International Congress of Phonetic Sciences (ICPhS)*, San Francisco, **3**, 1885-1888.

Fant, G. (1960). *Acoustic Theory of Speech Production*. The Hague, Netherlands: Mouton.

Fant, G. (1969). Stops in CV syllables. *Speech Transmission Labs Quarterly Progress Reports*, **41**. Reproduced in Baken, R. J. & Daniloff, R. G. (1991). *Readings in Clinical Spectrography of Speech*. San Diego; New Jersey: Singular Publishing Group and Kay Elemetrics. 179-215.

Farnetani, E. (1997). Coarticulation and connected speech processes. In W. J. Hardcastle & J. Laver (Eds), *The Handbook of Phonetic Sciences*. Oxford: Blackwell. 371-404.

Farnetani, E., Hardcastle, W. J. & Marchal, A. (1989). Cross-language investigation of lingual coarticulatory processes using EPG. In J. P. Tubach & J. J. Mariani (Eds), *Eurospeech 1989. European Conference on Speech Communication and Technology*. 429-432.

Farnetani, E. & Recasens, D. (1993). Anticipatory consonant-to-vowel coarticulation in the production of VCV sequences in Italian. *Language and Speech*, **36**, 279-302.

Fowler, C. A. (1977). *Timing Control in Speech Production*. Bloomington, IN: Indiana University Linguistics Club.

Fowler, C. A. (1980). Coarticulation and theories of extrinsic timing. *Journal of Phonetics*, **8**, 113-133.

Fowler, C. A. (2005). Parsing coarticulated speech in perception: effects of coarticulation resistance. *Journal of Phonetics*, **33**, 199-213.

Fowler, C. A. & Brancazio, L. (2000). Coarticulation resistance of American consonants and its effects on transconsonantal vowel-to-vowel coarticulation. *Language and Speech*, **43**, 1-41.

Fowler, C. A., Rubin, P., Remez, R. E. & Turvey, M. T. (1980). Implications for speech production of a general theory of action. In B. Butterworth (Ed.), *Language Production*. New York: Academic Press. 373-420.

Fowler, C. A. & Saltzman, E. (1993). Coordination and coarticulation in speech production. *Language and Speech*, **36**, 171-195.

Frisch, S. (2005). The phonetics of speech errors. [Oral paper, Ultrafest III 2005 (Ultrasound meeting), Arizona Phonological Imaging Lab, University of Arizona, Tucson, AZ, USA.]

Fromkin, V. (1968). Speculations on performance models. *Journal of Linguistics*, **4**, 47-68.

Fromkin, V. (Ed.) (1985). *Phonetic Linguistics: Essays in Honor of Peter Ladefoged*. Orlando: Academic Press.

Fromkin, V., Rodman, R. & Hyams, N. (2003). *An Introduction to Language*. 7th edn. Boston, MA: Thomson, Heinle.

Fuchs, S., Hoole, P., Brunner, J. & Inoue, M. (2004). The trough effect – an aerodynamic phenomenon? In J. Slifka, S. Manuel & M. Matthies (Eds), *Proceedings of the Conference "From Sound to Sense: 50+ Years of Discoveries in Speech Communication"*. Massachusetts Institute of Technology. C-25–C-30.

Fujimura, O. (2000). C/D model and prosodic control of articulatory behavior. *Phonetica*, **57**, 128-138.

Gay, T. (1974). Some electromyographic measures of coarticulation in VCV utterances. *Haskins Laboratories Status Report on Speech Research*, **SR-44**, 137-145.

Gay, T. (1977). Articulatory movements in VCV sequences. *Journal of the Acoustical Society of America*, **62**, 183-193.

Gay, T. (1978). Articulatory units: segments or syllables? In A. Bell & J. B. Hooper (Eds), *Syllables and Segments*. Amsterdam: North Holland. 121-131.

Gay, T. (1979). Coarticulation in some Consonant-Vowel and Consonant Cluster-Vowel Syllables. In B. Lindblom & G. Fant (Eds), *Frontiers of Speech Communication Research*. London: Academic Press. 69-76.

Gay, T., Ushijima, T., Hirose, H. & Cooper, F. S. (1974). Effect of speaking rate on labial consonant-vowel articulation. *Journal of Phonetics*, **2**, 46-63.

Gibbon, F. E. (1999). Undifferentiated lingual gestures in children with articulation/phonological disorders. *Journal of Speech, Language, and Hearing Research*, **42**, 382-397.

Gibbon, F. E. (2006). Current perspectives on clinical applications of EPG. [Oral presentation, 11th Meeting of the International Clinical Phonetics and Linguistics Association, Dubrovnik, Croatia.]

Gibbon, F. E., Hardcastle, W. J. & Nicolaidis, K. (1993). Temporal and spatial aspects of lingual coarticulation in /kl/ sequences: a cross-linguistic investigation. *Language and Speech*, **36**, 261-277.

Gibbon, F. & Wolters, M. (2005). A new application of ultrasound to image tongue behaviour in cleft palate speech. [Poster, The Craniofacial Society of Great Britain and Ireland Annual Scientific Conference 2005, Swansea, UK.]

References

Gick, B. (2002). The use of ultrasound for linguistic phonetic fieldwork. *Journal of the International Phonetic Association*, **32**, 113-122.

Gick, B. (2005). Situation-specific strategies for addressing head and transducer movement in ultrasound speech research. [Oral paper, Ultrafest III 2005 (Ultrasound meeting), Arizona Phonological Imaging Lab, University of Arizona, Tucson, AZ, USA.]

Gick, B., Bernhardt, B., Bacsfalvi, P. & Wilson, I. (forthcoming). Ultrasound imaging applications in second language acquisition. In J. Hansen & M. Zampini (Eds), *State-of-the-Art Issues in Second Language Phonology*. Cambridge: Cambridge University Press.

Gick, B., Bird, S. & Wilson, I. (2005). Techniques for field application of lingual ultrasound imaging. *Clinical Linguistics & Phonetics*, **19**, 503-514.

Gick, B., Campbell, F. & Oh, S. (2001). A cross-linguistic study of articulatory timing in liquids. *Journal of the Acoustical Society of America*, **110**, 2656.

Gick, B., Campbell, F., Oh, S. & Tamburri-Watt, L. (2006). Toward universals in the gestural organization of syllables: a cross-linguistic study of liquids. *Journal of Phonetics*, **34**, 49-72.

Gick, B. & Rahemtulla, S. (2004). Recent developments in quantitative analysis of ultrasound tongue data. [Oral paper, Ultrafest II 2004, Department of Linguistics, University of British Columbia, Vancouver, Canada.]

Gick, B. & Wilson, I. (2001). Pre-liquid excrescent schwa: what happens when vocalic targets conflict. In P. Dalsgaard, B. Linberg & H. Brenner (Eds), *Proceedings of the 7th European Conference on Speech Communicacation and Technology (Eurospeech 2001).* Aalborg, Denmark: Center for Personkommunikation (CPK). 273-276.

Gick, B. & Wilson, I. (forthcoming). Excrescent schwa and vowel laxing: cross-linguistic responses to conflicting articulatory targets. In L. M. Goldstein, D. H. Whalen, & C. T. Best (Eds), *Laboratory Phonology 8*. Berlin: Mouton de Gruyter.

Gick, B., Wilson, I., Koch, K. & Cook, C. (2004). Language-specific articulatory settings: evidence from inter-utterance rest position. *Phonetica*, **61**, 220-233.

Goldstein, L. & Fowler, C. (2003). Articulatory phonology: a phonology for public language use. In A. S. Meyer & N. O. Schiller (Eds), *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities*. Berlin: Mouton de Gruyter. 159-207.

Goldstein, L., Nam, H. & Saltzman, E. (2006). A coupled oscillator model of speech production planning. [Oral paper, 5th International Conference on Speech Motor Control, Radboud University Nijmegen, Nijmegen, Holland.]

Halle, M., Hughes, G. W. & Radley, I. P. A. (1957). Acoustic properties of stop consonants. *Journal of the Acoustical Society of America*, **29**, 107-116.

Hamann, S. (2004). Retroflex fricatives in Slavic languages. *Journal of the International Phonetic Association*, **34**, 53-67.

Hardcastle, W. J. (1976). *Physiology of Speech Production*. London: Academic Press.

Hardcastle, W., Gibbon, F. & Jones, W. (1991a). Visual display of tongue-palate contact: electropalatography in the assessment and remediation of speech disorders. *British Journal of Disorders of Communication*, **26**, 41-74.

Hardcastle, W., Gibbon, F. & Nicolaidis, K. (1991b). EPG data reduction methods and their implications for studies of lingual coarticulation. *Journal of Phonetics*, **19**, 251-266.

Hardcastle, W. & Hewlett, N. (1999). Introduction. In W. Hardcastle & N. Hewlett (Eds), *Coarticulation: Theory, Data and Techniques*. Cambridge: Cambridge University Press. 1-4.

Hardcastle, W., Jones, W., Knight, C., Trudgeon, A. & Calder, G. (1989). New developments in electropalatography: a state-of-the-art report. *Clinical Linguistics and Phonetics*, **3**, 1-38.

Hawkins, S. (1992). An introduction to task dynamics. In G. J. Docherty & D. R. Ladd (Eds), *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*. Cambridge: Cambridge University Press. 9-25.

Henke, W. (1966). *Dynamic Articulatory Model of Speech Production Using Computer Simulation.* PhD dissertation, MIT.

Hewlett, N. & Beck, J. (2006). *An Introduction to the Science of Phonetics*. Mahwah, N.J.; London: Lawrence Erlbaum.

Hewlett, N., Vazquez Alvarez, Y. & Zharkova, N. (2004). Ultrasound study of coarticulation & the “Trough Effect” in symmetrical VCV syllables: a report of work in progress. [Oral paper, Ultrafest II 2004, Department of Linguistics, University of British Columbia, Vancouver, Canada.]

Houde, R. A. (1967). *A Study of Tongue Motion During Selected Speech Sounds*. PhD dissertation, Speech Communication Research Laboratory, Santa Barbara.

Iskarous, K. (2005). Patterns of tongue movement. *Journal of Phonetics*, **33**, 363-381.

Jakobson, R., Fant, G. & Halle, M. (1951). *Preliminaries to Speech Analysis: the Distinctive Features and Their Corerlates*. Cambridge, MA: MIT Press.

Jones, W. J. & Hardcastle, W. J. (1995). New developments in EPG3 software. *European Journal of Disorders of Communication*, **30**, 183-192.

Karbownicki, L. (2004). *Investigation of the Coarticulation Effects on [h] When Preceding a Vowel.* BSc, Honours project. Queen Margaret University College, Edinburgh.

Kasatkin, L. L. (2006). *Sovremennyj russkij jazyk. Fonetika*. (*Modern Russian. Phonetics*.) Moscow: Academia.

Keating, P. A. (1988). Underspecification in phonetics. *Phonology*, **5.2**, 275-292.

Keating, P. A. (1990). The window model of coarticulation: articulatory evidence. In J. Kingston & M. E. Beckman (Eds), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. Cambridge: Cambridge University Press. 451-470.

Keating, P., Lindblom, B., Lubker, J. & Kreiman, J. (1994). Variability in jaw height for segments in English and Swedish VCVs. *Journal of Phonetics*, **22**, 407-422.

References

Keller, E. (1987). Ultrasound measurements of tongue dorsum movements in articulatory speech impairments. In J. H. Ryalls (Ed.), *Phonetic Approaches to Speech Production in Aphasia and Related Disorders*. San Diego, CA: College-Hill Press. 93-112.

Keller, E. & Ostry, D. (1983). Computerized measurement of tongue dorsum movement with pulsed echo ultrasound. *Journal of the Acoustical Society of America*, **73**, 1309-1315.

Kelsey, C. A., Woodhouse, R. J. & Minifie, F. D. (1969*).* Ultrasonic observations of coarticulation in the pharynx. *Journal of the Acoustical Society of America*, **46**, 1016-1018.

Kent, R. D. (1983). *The segmental organization of speech*. In P. F. MacNeilage (Ed.), *The Production of Speech*. New York: Springer Verlag. 57-89.

Kent, R. D. & Read, C. (2002). *The Acoustic Analysis of Speech*. 2nd edn. Albany, NY: Singular Thomson Learning.

Kodzasov, S. V. & Krivnova, O. F. (2001). *Obschaja fonetika*. (*General Phonetics*.) Moscow: Russian State University for the Humanities.

Kozhevnikov, V. A. & Chistovich, L. A. (Eds) (1965). *Speech: Articulation and Perception*. Washington: Joint Publications Research Service.

Kühnert, B. & Nolan, F. (1999). The origin of coarticulation. In W. Hardcastle & N. Hewlett (Eds), *Coarticulation: Theory, Data and Techniques*. Cambridge: Cambridge University Press. 7-30.

Kuznetsov, V. I. (1997). Vokalizm svyaznoj rechi. (*Vowels in Connected Speech*). St. Petersburg: State University of St. Petersburg.

Kuznetsov, V. I. (2000). Fonemnyj sostav slova v svyaznoj rechi (*Phonemic composition of words in connected speech*). In L. V. Bondarko (Ed.), *Fonologija rechevoj dejatelnosti*. (*Phonology of Speech Production*). St. Petersburg: State University of St. Petersburg. 47-73.

Ladefoged, P. (1957). Use of palatography. *Journal of Speech and Hearing Disorders*, **22**, 764-774.

Ladefoged, P. (1967). *Linguistic Phonetics. UCLA Working Papers in Phonetics 6*. Los Angeles: UCLA.

Ladefoged, P. (2001). *A Course in Phonetics*. 4th edn. Orlando, FL: Harcourt College Publishers.

Ladefoged, P. (2005). *Vowels and Consonants*. 2nd edn. Malden, MA: Blackwell Publishing.

Li, M., Kambhamettu, C. & Stone, M. (2003). EdgeTrak, a program for band-edge extraction and its applications. [Paper, Sixth IASTED International Conference o Computers, Graphics and Imaging, Honolulu, HI, USA.]

Li, M., Kambhamettu, C. & Stone, M. (2005). Automatic contour tracking in ultrasound images. *Clinical Linguistics and Phonetics*, **19**, 545-554.

Lindblom, B., Sussman, H. M., Modarresi, G. & Burlingame, E. (2002). The trough effect: implications for speech motor programming. *Phonetica*, **59**, 245-262.

Lubker, J. & Gay, T. (1982). Anticipatory labial coarticulation: experimental, biological and linguistic variables. *Journal of the Acoustical Society of America*, **71**, 437-448.

MacKay, I. R. A. (1977). Tenseness in vowels: an ultrasonic study. *Phonetica*, **34**, 325-351.

Matusevich, M. I. (1948). *Vvedenije v obschuju fonetiku*. (*Introduction to General Phonetics*.) 2nd edn. Leningrad: Uchpedgiz.

McAllister, R. (1978). Temporal asymmetry in labial coarticulation. *Papers from the Institute of Linguistics, University of Stokholm*, **35**, 1-29.

McAllister, R. & Engstrand, O. (1991). Some cross-language aspects of coarticulation. *Actes du XIIème Congrès International des Sciences Phonétiques*, Aix-en-Provence, **5**, 18-21.

McAllister, R. & Engstrand, O. (1992). Interpretations of tongue movement patterns in VCV sequences. In D. Huber (Ed.), *Papers from the Sixth Swedish Phonetics Conference held in Gothenburg, May 20-22, Technical Report No. 10, Department of Information Theory, School of Electrical and Computer Engineering, Chalmers University of Technology*, Gothenburg, 115-120.

Ménard, L., Aubin, J., Roy, J.-P. & Leclerc, A. (2005). Acoustic, aerodynamic and articulatory correlates of contrastive focus in French. [Oral paper, Ultrafest III 2005 (Ultrasound meeting), Arizona Phonological Imaging Lab, University of Arizona, Tucson, AZ, USA.]

Merlin Ultrasound Scanner Type 1101, User Guide (2000). Gentofte: B-K Medical.

Minifie, F., Kelsey, G., Zagzebski, J. & King, T. (1971). Ultrasonic scans of the dorsal surface of the tongue. *Journal of the Acoustical Society of America*, **49**, 1857-1860.

Modarresi, G., Sussman, H., Lindblom, B. & Burlingame, E. (2004). An acoustic analysis of the bidirectionality of coarticulation in VCV utterances. *Journal of Phonetics*, **32**, 291-312.

Moll, K. L. & Daniloff, R. G. (1971). Investigation of the timing of velar movements during speech. *Journal of the Acoustical Society of America*, **50**, 678-684.

Morrish, K. A., Stone, M., Sonies, B. C., Kurtz, D. & Shawker, T. (1984). Characterization of tongue shape. *Ultrasonic Imaging*, **6**, 37-47.

Morrish, K. A., Stone, M., Shawker, T. H. & Sonies, B. C. (1985). Distinguishability of tongue shape during vowel production. *Journal of Phonetics*, **13**, 189-203.

Muhnall, K. G. & Ostry, D. J. (1985). Ultrasonic measurement of laryngeal kinematics. In I. R. Titze & R. C. Scherer (Eds), *Vocal Fold Physiology: Biomechanics, Acoustics and Phonatory Control*. Denver: Denver Center for the Performing Arts. 145-162.

Nguyen, N., Marchal, A. & Content, A. (1996). Modeling tongue-palate contact patterns in the production of speech. *Journal of Phonetics*, **24**, 77-97.

Niimi, S. & Simada, Z. (1980). Ultrasonic investigation of tongue shape. *Japanese Journal of Logopedics and Phoniatrics*, **21**, 121-125.

Oh, S. (2002). Cross-language blending of /l/ gestures by Korean-English bilingual children". *26th Boston University Conference on Language Development (BUCLD)*, **2**, 473-484.

Öhman, S. E. G. (1966). Coarticulation in VCV utterances: spectrographic measurements. *Journal of the Acoustical Society of America*, **39**, 151-168.

Öhman, S. E. G. (1967). Numerical model of coarticulation. *Journal of the Acoustical Society of America*, **41**, 310-319.

References

Ostry, D. J., Gribble, P. L. & Gracco, V. L. (1996). Coarticulation of jaw movements in speech production: is context sensitivity in speech kinematics centrally planned? *The Journal of Neuroscience*, **16**, 1570-1579.

Parthasarathy, V., Stone, M. & Prince, J. L. (2005). Spatiotemporal Visualization of the Tongue Surface using Ultrasound and Kriging (SURFACES). *Clinical Linguistics and Phonetics*, **19**, 529-544.

Parush, A., Ostry, D. J. & Munhall, K. G. (1983). A kinematic study of lingual articulation in VCV sequences. *Journal of the Acoustical Society of America*, **74**, 1115-1125.

Perkell, J. S. (1969). *Physiology of Speech Production: Results and Implications of a Quantitative Cineradiographic Study*. Cambridge, MA: MIT Press.

Perkell, J. S. (1986). Coarticulation strategies: preliminary implications of a detailed analysis of lower lip protrusion movements. *Speech Communication*, **5**, 47-68.

Perkell, J. S. & Matthies, M. L. (1992). Temporal measures of anticipatory labial coarticulation for the vowel /u/: within- and cross-subject variability. *Journal of the Acoustical Society of America*, **91**, 2911-2925.

Pierrehumbert, J. & Talkin, D. (1992). Lenition of [h] and glottal stop. In J. Docherty & D. R. Ladd (Eds), *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*. Cambridge: Cambridge University Press. 90-117.

Recasens, D. (1984). V-to-C coarticulation in Catalan VCV sequences: an articulatory and acoustical study. *Journal of Phonetics*, **12**, 61-73.

Recasens, D. (1987). An acoustic analysis of V-to-C and V-to-V coarticulatory effects in Catalan and Spanish VCV sequences. *Journal of Phonetics*, **15**, 299-312.

Recasens, D. (1999). Lingual coarticulation. In W. Hardcastle & N. Hewlett (Eds), *Coarticulation: Theory, Data and Techniques*. Cambridge: Cambridge University Press. 80-104.

Recasens, D. (2002a). An EMA study of VCV coarticulatory direction. *Journal of the Acoustical Society of America*, **111**, 2828-2841.

Recasens, D. (2002b). Patterns of jaw coarticulatory direction and linguomandibular coordination in VCV sequences. *ZAS Papers in linguistics*, **28**, 83-92.

Recasens, D. (2002c). Production mechanisms in consonant clusters and phonological implications. [Paper, LabPhon 8, Yale University & Haskins Laboratory, New Haven, CT, USA.]

Recasens, D. (2004). The effect of syllable position on consonant reduction (evidence from Catalan consonant clusters). *Journal of Phonetics*, **32**, 435-453.

Recasens, D. & Pallarès, M. D. (1999). A study of /ɾ/ and /r/ in the light of the “DAC” coarticulation model. *Journal of Phonetics*, **27**, 143-169.

Recasens, D. & Pallarès, M. D. (2000). A study of F1 coarticulation in VCV sequences. *Journal of Speech and Hearing Research*, **43**, 501-512.

Recasens, D. & Pallarès, M. D. (2001). Coarticulation, assimilation and blending in Catalan consonant clusters. *Journal of Phonetics*, **29**, 273-301.

Recasens, D., Pallarès, M. D. & Fontdevila, J. (1997). A model of lingual coarticulation based on articulatory constraints. *Journal of the Acoustical Society of America*, **102**, 544-561.

Recasens, D., Pallarès, M. D. & Fontdevila, J. (1998). An electropalatographic and acoustic study of temporal coarticulation for Catalan dark /l/ and German clear /l/. *Phonetica*, **55**, 53-79.

Reformatskij, A. A. (Ed.) (1970). Iz istorii otechestvennoj fonologii (*From the History of the Russian Phonology*). Moscow: Nauka.

Reformatskij, A. A. (2006). Vvedenije v jazykovedenije. (*Introduction to Linguistics*). 5th edn. Moscow: Aspect Press.

Saltzman, E. L. (1991). The task dynamic model in speech production. In H. F. M. Peters, W. Hulstijn & C. W. Starkweather (Eds), *Speech Motor Control and Stuttering*. Amsterdam: Elsevier Science Publishers. 37-52.

Saltzman, E. & Byrd, D. (2000). Task-dynamics of gestural timing: phase windows and multifrequency rhythms. *Human Movement Science*, **19**, 499-526.

Saltzman, E. & Kelso, J.-A. S. (1987). Skilled actions: a task dynamic approach. *Psychological Review,* **94***,* 84-106.

Saltzman, E. L. & Munhall, K. G. (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology*, **1**, 333-382.

Scherba, L. V. (1912). *Russkije glasnyje v kachestvennom i kolichestvennom otnoshenii*. (*Russian Vowels in Qualitative and Quantitative Respect*.) St. Petersburg: J. N. Erlich Press.

Scobbie, J. & Sebregts, K. (2005). Getting at variation with ultrasound: examples from Scottish and Dutch /r/. [Oral paper, Ultrafest III 2005 (Ultrasound meeting), Arizona Phonological Imaging Lab, University of Arizona, Tucson, AZ, USA.]

Scobbie, J. & Stuart-Smith, J. (2006). Where the [l] [r] they? An ultrasound tongue imaging study of Scottish English liquids. [Oral paper, British Association of Academic Phoneticians Colloquium 2006, Queen Margaret University College, Edinburgh, UK.]

Scobbie, J., Wood, S. & Wrench, A. (2004). Advances in EPG for treatment and research: an illustrative case study. *Clinical Linguistics and Phonetics*, **18,** 373-389.

Shawker, T. H., Sonies, B. C. & Stone, M. (1984). Soft tissue anatomy of the tongue and floor of the mouth: an ultrasound demonstration*.* *Brain and Language*, **21**, 335-350.

Shawker, T. H., Stone, M. & Sonies, B. C. (1985). Tongue pellet tracking by ultrasound: development of a reverberation pellet. *Journal of Phonetics*, **13**, 135-146.

Skolnick M. L., Zagzebski, J. A. & Watkin, K. L. (1975). Two dimensional ultrasonic demonstration of lateral pharyngel wall movement in real time – a preliminary report. *Cleft Palate Journal*, **12**, 299-303.

Sonies, B. C., Shawker, T. H., Hall, T. E., Gerber, L. H. & Leighton, S. B. (1981). Ultrasonic visualization of tongue motion during speech. *Journal of the Acoustical Society of America*, **70**, 683-686.

References

Stevens, K. N. (1971). Airflow and turbulence noise for fricative and stop consonants: static considerations. *Journal of the Acoustical Society of America*, **50**, 1180-1192.

Stone, M. (1990). A three-dimensional model of tongue movement based on ultrasound and x-ray microbeam data. *Journal of the Acoustical Society of America*, **87**, 2207-2217.

Stone, M. (1991). Toward a model of three-dimensional tongue movement. *Journal of Phonetics*, **19**, 309-320.

Stone, M. (1995). How the tongue takes advantage of the palate during speech. In F. Bell-Berti & J. R. Laurence (Eds), *Producing Speech: Contemporary Issues for Katherine Safford Harris*. New York: AIP Press. 143-153.

Stone, M. (1999). Laboratory techniques for investigating speech articulation. In W. J. Hardcastle & J. Laver (Eds), *The Handbook of Phonetic Sciences*. Oxford: Blackwell. 11-32.

Stone, M. (2005). A guide to analyzing tongue motion from ultrasound images. *Clinical Linguistics and Phonetics*, **19**, 455-502.

Stone, M. & Davis, E. P. (1995). A head and transducer support system for making ultrasound images of tongue/jaw movement. *Journal of Acoustical Society of America*, **6**, 3107-3112.

Stone, M., Epstein, M. & Iskarous, K. (2004). Functional segments in tongue movement. *Clinical Linguistics and Phonetics*, **18**, 507-521.

Stone, M., Faber, A., Raphael, L. J. & Shawker, T. H. (1992). Cross-sectional tongue shape and linguopalatal contact patterns in [s], [ʃ], and [l]. *Journal of Phonetics*, **20**, 253-270.

Stone, M. & Lele, S. (1992). Representing the tongue surface with curve fits. *Proceedings of the International Conference on Spoken Language Processing (ICSLP 92)*. Banff, Canada. 875-878.

Stone, M. & Lundberg, A. (1996). Three-dimensional tongue surface shapes of English consonants and vowels. *Journal of the Acoustical Society of America*, **99**, 3728-3737.

Stone, M., Morrish, K. A., Sonies, B. C. & Shawker, T. H. (1987). Tongue curvature: a model of shape during vowel production. *Folia phoniatrica*, **39**, 302-315.

Stone, M., Shawker, T. H., Talbot, T. & Rich, A. (1988). Cross-sectional tongue shape during the production of vowels. *Journal of the Acoustical Society of America*, **83**, 1586-1596.

Stone, M., Sonies, B. C., Shawker, T. H., Weiss, G. & Nadel, L. (1983). Analysis of real-time ultrasound images of tongue configuration using a grid-digitizing system. *Journal of Phonetics*, **11**, 207-218.

Stone, M. & Vatikiotis-Bateson, E. (1995). Trade-offs in tongue, jaw, and palate contributions to speech production. *Journal of Phonetics*, **23**, 81-100.

Su, L., Li, K. P. & Fu, K. S. (1974). Identification of speakers by use of nasal coarticulation. *Journal of the Acoustical Society of America*, **56**, 1876-1882.

Sussman, H. M. & Westbury, J. R. (1981). The effects of antagonistic gestures on temporal and amplitude parameters of anticipatory labial coarticulation. *Journal of Speech and Hearing Research*, **46**, 16-24.

Svirsky, M., Stevens, K., Matthies, M., Manzella, J., Perkell, J. & Wilhelms-Tricarico, R. (1997). Tongue surface displacement during bilabial stops. *Journal of the Acoustical Society of America*, **102**, 562-571.

Trubetzkoy, N. S. (1939). *Gründzuge der Phonologie*. Travaux du Cercle Linguistique de Prague VII. Prague.

Unser, M. & Stone, M. (1992). Automated detection of the tongue surface in sequences of ultrasound images. *Journal of the Acoustical Society of America*, **91**, 3001-3007.

Vazquez Alvarez, Y. (2006). Cross-linguistic study of closure-related tongue perturbations. [Oral paper, Speech and Hearing Sciences Research Seminar, Queen Margaret University College, Edinburgh.]

Vazquez Alvarez, Y., Hewlett, N. & Zharkova, N. (2004). An ultrasound study of the “Trough Effect”. [Poster, British Association of Academic Phoneticians Colloquium 2004, University of Cambridge, Cambridge, UK.]

Vilain, A., Abry, C. & Badin, P. (1999). Motor equivalence evidenced by articulatory modelling. *Proceedings of Eurospeech 1999*, *Budapest*. 169-172.

Whalen, D. H. & Gick, B. (2001). Intrinsic F_0 and pharyngeal width in ATR languages. *Journal of the Acoustical Society of America*, **110**, 2761.

Wilson, I. (2005). Articulatory settings of French-English bilingual speakers. *Journal of the Acoustical Society of America*, **117**, 2426.

Wilson, I. (2006). *Articulatory Settings of French and English Monolingual and Bilingual Speakers.* PhD dissertation, University of British Columbia.

Wilson, I. (forthcoming). The effects of post-velar consonants on vowels in Nuuchahnulth: auditory, acoustic, and articulatory evidence. *Canadian Journal of Linguistics.*

Wood, S. (1997). *An Electropalatographic Study of Speech Sound Errors in Adults with Acquired Aphasia*. PhD dissertation, Open University/Queen Margaret College.

Wrench, A. A. (2006). *Articulate Assistant User Guide version 1.12.* Retrieved May 2, 2006, from http://www.articulateinstruments.com.

Wrench, A. A. & Scobbie, J. M. (2003). Categorising vocalisation of English /l/ using EPG, EMA and ultrasound. In S. Palethorpe & M. Tabain (Eds), *Proceedings of the 6th International Seminar on Speech Production (ISSP Sydney)*. 314-319.

Xu, Y. & Liu, F. (forthcoming). Tonal alignment, syllable structure and coarticulation: toward an integrated model. *Italian Journal of Linguistics*.

Yuen, I., Lee, A. & Gibbon, F. (2005). EPG study of bilabials in intervocalic contexts. [Oral paper, 4th International EPG Symposium, Queen Margaret University College, Edinburgh, UK.]

Zagzebski, J. A. (1975). Ultrasonic measurement of lateral pharyngeal wall motion at two levels in the vocal tract. *Journal of Speech and Hearing Research*, **18**, 308-318.

Zharkova, N. & Hewlett, N. (2006). Studying coarticulation resistance with ultrasound. [Poster, 5th International Conference on Speech Motor Control, Radboud University Nijmegen, Nijmegen, The Netherlands.]

References

Zinder, L. R. (1979). *Obschaja fonetika*. (*General Phonetics*.) 2nd edn. Moscow: Vysshaja shkola.

Zinder, L. R., Bondarko, L. V. & Verbitskaja, L. A. (1964). Akusticheskaja harakteristika razlichija tverdyh i myagkih soglasnyh v russkom jazyke. (*Acoustic characteristics of hard and palatalised consonants in Russian*.) *Uchenyje Zapiski Leningradskogo Gosudarstvennogo Universiteta*, **325**, 28-36.

Zinder, L. R., Bondarko, L. V. & Verbitskaja, L. A., Svetozarova, N. D. (1968). Nekotoryje priznaki chlenenija russkoj rechi na slova. (*Some characteristics of parsing Russian speech into words*.) [Paper, 6th Acoustic Conference of the USSR, Moscow, USSR.]

APPENDIX I. SUBJECT RECRUITMENT DOCUMENTS

APPENDIX I-1. Experiment 1

Request for native Russian-speaking volunteers

I am looking for students/staff who would be willing to participate in an ultrasound study. The aim of this study is to investigate tongue movements during the pronunciation of vowel-consonant-vowel syllables.

The ultrasound technique used is non-invasive and is subject to rigorous safety assessments. There are no known risks associated with this technique.

Participants will be native Russian speakers who do not suffer from any speech disorder. I attach some information about what is involved.

Length of the experiment	30 minutes
Location	Queen Margaret University College, Edinburgh Speech and Language Sciences Room 113, soundproofed room
Contact	Natalia Zharkova (PhD student). Tel: 0131 317 ex. 3687 e-mail: NZharkova@QMUC.ac.uk

Thank you for your interest,

Natalia Zharkova,
Queen Margaret University College

PS. If you are personally known to me, you will receive this request via my supervisor, Dr. Nigel Hewlett.

Поиск добровольцев - носителей русского языка

Я ищу студентов/сотрудников, которые хотели бы принять участие в фонетическом эксперименте с использованием ультразвука. Цель исследования – изучение движения языка во время произнесения слогов модели Гласный-Согласный-Гласный.

Используемая методика – ультразвук – неинвазивная и безопасная для испытуемых. О возможностях риска, связанного с ультразвуком, науке не известно.

Участники должны быть носителями русского языка, не страдающими речевыми расстройствами. Более подробная информация об эксперименте прилагается ниже.

Продолжительность эксперимента	30 минут
Место	Queen Margaret University College, Edinburgh Speech and Language Sciences Room 113, заглушенная комната
Контактное лицо	Natalia Zharkova (PhD student). Tel: 0131 317 ex. 3687 e-mail: NZharkova@QMUC.ac.uk

Спасибо за проявленный интерес,

Наталья Жаркова

Natalia Zharkova,
Queen Margaret University College

PS. Если Вы лично знакомы со мной, Вы получите эту информацию через моего научного руководителя (Dr. Nigel Hewlett).

Queen Margaret University College
EDINBURGH

The Application of Ultrasound in the Investigation of Speech Dynamics

Information sheet for participants

This study uses ultrasound to investigate tongue movements during the pronunciation of vowel-consonant-vowel syllables.

The ultrasound technique used is non-invasive and is subject to rigorous safety assessments. At all levels of intensity used for diagnostic imaging, there are no known risks associated with ultrasound and there are no specific dangers or safety requirements.

You will be asked to sit next to an ultrasound scanner in a soundproofed studio. You will use a helmet, which will ensure that the ultrasound transducer can be correctly positioned beneath your chin. The end of the transducer will be covered in medical gel, or rest on a ¾ in. acoustic standoff pad. You will be asked to put some jelly (ordinary food jelly suitable for vegetarians) into your mouth, before you read the phrases, to facilitate imaging the hard palate (you will choose whether to swallow the jelly or to spit it out before the recordings). You will be filmed during the experiment.

Your task will be to read a list of sentences displayed on a computer screen.

The whole procedure should not take longer than 30 minutes.

The session will be recorded for later acoustic analysis. The measurements will be tested for statistical significance.

All data will be anonymised. There is a possibility that someone known to you might recognise your voice, if it was included in a presentation. If a section of video is played in a presentation you will be recognizable by that. You will not be mentioned by name in any report or presentation.

You are free to withdraw from the study at any stage without giving a reason.

If you would like to consult an independent person, who knows about this project but is not involved in it, you are welcome to contact Dr. Alan Wrench, ex. 3423, Email: **awrench@qmuc.ac.uk**.

Natalia Zharkova
PhD student, Speech and Language Sciences
Queen Margaret University College, Edinburgh

Queen Margaret University College
EDINBURGH

Использование ультразвука в изучении речевой динамики

Информация для участников

В данной работе с помощью ультразвука исследуются движения языка во время произнесения слогов модели Гласный-Согласный-Гласный.

Используемая методика – ультразвук – неинвазивная и безопасная для испытуемых. На всех уровнях интенсивности, используемых для ультразвукового сканирования, не существует риска, связанного с ультразвуком. Не требуются и особые меры предосторожности.

Вам предстоит сидеть рядом с ультразвуковым аппаратом в заглушенной комнате. На вас будет надет шлем, с помощью которого ультразвуковой сканер будет необходимым образом расположен под Вашим подбородком. Сканер будет смазан медицинским гелем, либо будет помещен на специальной подушечке. Вам необходимо будет взять в рот немного желе (обычного пищевого желе, подходящего для вегетарианцев), для того, чтобы можно было увидеть твердое нёбо (по Вашему выбору Вы сможете проглотить желе или выплюнуть его перед записью). Вас будут снимать на видеокамеру во время эксперимента.

Ваша задача будет состоять в том, чтобы читать предложения, появляющиеся на экране компьютера.

Вся процедура занимает не более 30 минут.

Будет производиться запись для последующего акустического анализа. Статистическая достоверность измерений будет впоследствии специально проверена.

Все данные будут представлены анонимно. Есть возможность, что Ваши знакомые узнают Ваш голос, если эти данные будут включены в научный доклад. Ваше имя не будет упомянуто или эксплицитно представлено в отчетах или докладах.

Вы можете прервать свое участие в эксперименте в любой момент, без объяснения причин.

Если Вы хотели бы поговорить с человеком, который знает о данном исследовании, но не участвует в нем, обращайтесь к Dr. Alan Wrench, ex. 3423, Email: **awrench@qmuc.ac.uk**.

Наталья Жаркова (Natalia Zharkova)
PhD student, Speech and Language Sciences
Queen Margaret University College, Edinburgh

Queen Margaret University College
EDINBURGH

The Application of Ultrasound in the Investigation of Speech Dynamics

Consent Form

I have read and understood the subject information sheet and this consent form. I have had an opportunity to ask questions about my participation.

I understand that I am under no obligation to take part in this study.

I understand that I have the right to withdraw from this study at any stage without giving any reason, and that withdrawal would not affect my current or future studies.

I agree to participate in this study and to have my audio and visual data used for analysis, reports and presentations.

Name of participant: ____

Signature of participant: ____

Signature of investigator: _______________________________________

Date: __________________________

Further information is available from:

Natalia Zharkova NZharkova@QMUC.ac.uk
Speech and Language Sciences
Queen Margaret University College

Queen Margaret University College
EDINBURGH

Использование ультразвука в изучении речевой динамики

Форма согласия

Я прочел (прочла) и понял (поняла) информацию для участников и данную форму согласия. У меня была возможность задать вопросы, касающиеся моего участия в эксперименте.

Я понимаю, что не обязан (не обязана) участвовать в данном исследовании.

Я понимаю, что у меня есть право прервать мое участие в эксперименте в любой момент, без объяснения причин, и что это никак не повлияет на мои текущие или будущие занятия в университете.

Я согласен (согласна) участвовать в данном исследовании, а также согласен (согласна) на то, что мои аудио- и видео-данные будут использованы для анализа, отчетов и докладов.

Имя участника: ____

Подпись участника: ____

Подпись исследователя: ____

Дата: __________________________

За более подробной информацией обращайтесь к Наталье Жарковой:

Natalia Zharkova NZharkova@QMUC.ac.uk
Speech and Language Sciences
Queen Margaret University College

APPENDIX I-2. Experiment 2

Request for native English-speaking volunteers

(An ultrasound study of tongue movements in nonsense syllables)

I am looking for students/staff who would be willing to participate in an ultrasound study. The aim of this study is to investigate tongue movements during the pronunciation of some nonsense syllables.

The ultrasound technique used is non-invasive and is subject to rigorous safety assessments. There are no known risks associated with this technique.

Participants will be native British English speakers who do not suffer from any speech disorder. I attach some information about what is involved (see Information Sheet for Participants).

Length of the experiment	30 minutes
Location	Queen Margaret University College, Edinburgh Speech and Language Sciences Room 116a, Ultrasound Recording Studio

Contact Natalia Zharkova (PhD student).
Tel: 0131 317 ex. 3687
e-mail: NZharkova@QMUC.ac.uk

Thank you for your interest,

Natalia Zharkova,
Queen Margaret University College

PS. If you are personally known to me, you will receive this request via my supervisor, Dr. Nigel Hewlett.

Queen Margaret University College
EDINBURGH

An ultrasound study of tongue movements in nonsense syllables

Information sheet for participants

This study uses ultrasound to investigate tongue movements during the pronunciation of some nonsense syllables.

This ultrasound technique is subject to rigorous safety assessments. At all levels of intensity used for diagnostic imaging, there are no known risks associated with ultrasound and there are no specific dangers or safety requirements.

You will be asked to sit next to an ultrasound scanner in a sound-treated studio. You will use a helmet, which will ensure that the ultrasound transducer can be correctly positioned beneath your chin. The end of the transducer will be covered in medical gel, or rest on a ¾ in. acoustic standoff pad.

Your task will be to read a list of sentences displayed on a computer screen.

The whole procedure should not take longer than 30 minutes.

The session will be recorded for later acoustic and tongue contour analysis.

All data will be anonymised. You will not be mentioned by name in any report or presentation. However, if some of your data were played at a verbal presentation, there is the possibility that you may be recognizable by your voice.

You are free to withdraw from the study at any stage without giving a reason.

If you would like to consult an independent person, who knows about this project but is not involved in it, you are welcome to contact Dr. Alan Wrench, ex. 3423, Email: **awrench@qmuc.ac.uk**.

Natalia Zharkova
PhD student, Speech and Language Sciences
Queen Margaret University College, Edinburgh

Queen Margaret University College
EDINBURGH

An ultrasound study of tongue movements in nonsense syllables

Consent Form

I have read and understood the subject information sheet and this consent form. I have had an opportunity to ask questions about my participation.

I understand that I am under no obligation to take part in this study.

I understand that I have the right to withdraw from this study at any stage without giving any reason, and that withdrawal would not affect my current or future studies.

I agree to participate in this study and to have my audio and visual data used for analysis, reports and presentations.

Name of participant: ____

Signature of participant: ____

Signature of investigator: ____

Date: __________________________

Further information is available from:

Natalia Zharkova NZharkova@QMUC.ac.uk
Speech and Language Sciences
Queen Margaret University College

APPENDIX I-3. Experiment 3

Request for native English-speaking volunteers

(An ultrasound and EPG study of tongue movements in speech)

I am looking for students/staff who would be willing to participate in an ultrasound and EPG study. The aim of this study is to investigate tongue movements during the pronunciation of some particular sound sequences in speech.

Both ultrasound and EPG techniques used are subject to rigorous safety assessments. There are no known risks associated with these techniques.

Participants will be native British English speakers who do not suffer from any speech disorder, and have artificial EPG palates. I attach some information about what is involved (see Information Sheet for Participants).

Length of the experiment	30 minutes
Location	Queen Margaret University College, Edinburgh Speech and Language Sciences Room 113b, Ultrasound Recording Studio

Contact Natalia Zharkova (PhD student).
Tel: 0131 317 ex. 3687
e-mail: NZharkova@QMUC.ac.uk

Thank you for your interest,

Natalia Zharkova,
Queen Margaret University College

PS. If you are personally known to me, you will receive this request via my supervisor, Dr. Nigel Hewlett.

Queen Margaret University College
EDINBURGH

An ultrasound and EPG study of tongue movements in speech

Information sheet for participants

This study uses ultrasound and electropalatography to investigate tongue movements during the pronunciation of some particular sound sequences in speech.

Both ultrasound and EPG techniques used are subject to rigorous safety assessments. At all levels of intensity used for diagnostic imaging, there are no known risks associated with ultrasound and there are no specific dangers or safety requirements. There are no known risks associated with EPG technique, either.

You will be asked to sit next to an ultrasound scanner in a sound-treated studio. You will wear your artificial EPG palate. You will use a helmet, which will ensure that the ultrasound transducer can be correctly positioned beneath your chin. The end of the transducer will be covered in medical gel, or rest on a ¾ in. acoustic standoff pad.

Your task will be to read a list of sentences displayed on a computer screen.

The whole procedure should not take longer than 30 minutes.

The session will be recorded for later acoustic and tongue contour analysis.

All data will be anonymised. You will not be mentioned by name in any report or presentation. However, if some of your data were played at a verbal presentation, there is the possibility that you may be recognizable by your voice.

You are free to withdraw from the study at any stage without giving a reason.

If you would like to consult an independent person, who knows about this project but is not involved in it, you are welcome to contact Dr. Alan Wrench, ex. 3423, Email: **awrench@qmuc.ac.uk**.

Natalia Zharkova
PhD student, Speech and Language Sciences
Queen Margaret University College, Edinburgh

Queen Margaret University College
EDINBURGH

An ultrasound and EPG study of tongue movements in speech

Consent Form

I have read and understood the subject information sheet and this consent form. I have had an opportunity to ask questions about my participation.

I understand that I am under no obligation to take part in this study.

I understand that I have the right to withdraw from this study at any stage without giving any reason, and that withdrawal would not affect my current or future studies.

I agree to participate in this study and to have my audio and visual data used for analysis, reports and presentations.

Name of participant: ____

Signature of participant: ____

Signature of investigator: ____

Date: ___________________________

Further information is available from:

Natalia Zharkova NZharkova@QMUC.ac.uk
Speech and Language Sciences
Queen Margaret University College

APPENDIX II.

Example of a text file with xy data exported from the ultrasound analysis software (Articulate Assistant).

The first line contains initials of the subject, the target carrier phrase as it appears on the screen for the subject, and the abbreviation for the annotation used in the analysis process ("/i/v1" stands for the V1 of the /ipʲi/ sequence). The first column of the second line contains the x value of the first data point on the V1 curve, in millimetres (76.17829). The second column of the second line contains the y value of the first data point on the V1 curve, in millimetres (43.16770). Columns 3 and 4 of the second line contain x and y values of the second data point on the V1 curve, respectively. Column 5 of the second line and Column 1 of the third line contain x and y values of the third data point on the V1 curve, respectively; and so on. There are 85 data points in the V1 curve.

The data for the C curve and the V2 curve are organised in the same way as the data for the V1 curve. The C curve is indicated by the annotation "/p/"; the V2 curve is indicated by the annotation "/i/v2". There are 85 data points in the C curve, and 87 data points in the V2 curve.

```
IG      Тогда Ира должна сказать иПИ семь тысяч раз, как задаток      /i/v1
        76.17829    43.16770    77.20600    43.78696    78.19967
        44.44101    79.15929    45.12984    80.08488    45.85347
        80.78086    46.46641    81.43106    47.12574    82.04109
        47.82581    82.61658    48.56097    83.16315    49.32556
        83.68643    50.11394    84.19203    50.92044    84.68557
        51.73941    85.17268    52.56519    85.65899    53.39214
        86.15011    54.21459    86.65166    55.02689    87.16927
        55.82339    87.70857    56.59844    88.27516    57.34637
        88.87468    58.06154    89.44300    58.71027    90.01864
        59.38069    90.60158    60.06985    91.19180    60.77478
        91.78929    61.49253    92.39402    62.22015    93.00598
        62.95466    93.62516    63.69312    94.25153    64.43257
        94.88508    65.17004    95.52579    65.90259    96.17365
        66.62725    96.82864    67.34106    97.49073    68.04107
        98.15993    68.72431    98.83619    69.38784    99.51952
        70.02869    100.2098    70.64390    100.9072    71.23052
        101.6116    71.78559    102.3230    72.30615    103.0414
        72.78924    103.7667    73.23191    104.4990    73.63119
        105.2382    73.98413    105.9843    74.28777    106.7373
        74.53915    107.4972    74.73531    108.2639    74.87330
        109.0374    74.95016    109.8178    74.96293    110.6050
        74.90866    111.4820    74.79335    112.3961    74.63735
        113.3298    74.43822    114.2659    74.19356    115.1869
        73.90095    116.0755    73.55796    116.9142    73.16219
```

117.6857	72.71121	118.1959	72.34799	118.6839
71.94375	119.1504	71.50195	119.5959	71.02607
120.0211	70.51959	120.4267	69.98597	120.8133
69.42868	121.1816	68.85121	121.5322	68.25702
121.8657	67.64959	122.1829	67.03239	122.4843
66.40890	122.7706	65.78258	123.0425	65.15690
123.3006	64.53536	123.5455	63.92140	123.8656
62.88885	124.1099	61.81972	124.3772	60.76374
124.7663	59.77066	125.3421	58.77340	126.0224
57.83058	126.8071	56.94220	127.6963	56.10824

IG Тогда Ира должна сказать иПИ семь тысяч раз, как задаток /p/

78.37575	43.90018	78.88880	44.27370	79.39479
44.65386	79.89371	45.04066	80.38556	45.43411
80.87035	45.83419	81.34808	46.24091	81.81873
46.65427	82.28232	47.07428	83.02385	47.80688
83.71708	48.58706	84.36861	49.40812	84.98505
50.26335	85.57299	51.14602	86.13905	52.04942
86.68983	52.96686	87.23193	53.89160	87.77195
54.81694	88.31651	55.73618	88.87220	56.64258
89.44562	57.52945	90.04339	58.39007	90.67210
59.21773	91.33836	60.00571	92.04878	60.74731
92.58902	61.27408	93.14335	61.81376	93.71117
62.36408	94.29186	62.92277	94.88481	63.48754
95.48940	64.05611	96.10503	64.62621	96.73108
65.19556	97.36694	65.76189	98.01200	66.32290
98.66564	66.87633	99.32725	67.41989	99.99623
67.95131	100.6719	68.46831	101.3538	68.96861
102.0411	69.44994	102.7334	69.91001	103.4300
70.34654	104.1303	70.75727	104.8336	71.13990
105.5395	71.49217	106.2471	71.81179	106.9560
72.09648	107.6655	72.34397	108.3751	72.55198
109.0840	72.71823	109.7917	72.84045	110.4976
72.91634	111.2011	72.94365	111.9015	72.92008
112.5983	72.84336	113.2908	72.71121	114.0635
72.51509	114.8620	72.27844	115.6720	72.00158
116.4792	71.68483	117.2694	71.32853	118.0282
70.93297	118.7414	70.49850	119.3948	70.02543
119.8496	69.62348	120.2642	69.17869	120.6430
68.69560	120.9903	68.17873	121.3105	67.63261
121.6078	67.06179	121.8865	66.47078	122.1512
65.86413	122.4059	65.24636	122.6552	64.62201
122.9033	63.99561	123.1546	63.37169	123.4134
62.75479	123.6840	62.14943	123.9707	61.56014
124.2780	60.99147	124.8793	59.94866	125.4867
58.90927	126.1001	57.87329	126.7196	56.84073

IG Тогда Ира должна сказать иПИ семь тысяч раз, как задаток /i/v2

76.17829	44.38851	77.04041	44.62079	77.88742
44.87683	78.75570	45.19139	79.58447	45.56549

```
80.37607    45.99573    81.13282    46.47869    81.85708
47.01097    82.55118    47.58917    83.21744    48.20988
83.85822    48.86970    84.47584    49.56522    85.07264
50.29304    85.65096    51.04974    86.21314    51.83194
86.76151    52.63621    87.29841    53.45916    87.82617
54.29738    88.34713    55.14747    88.86363    56.00601
89.37800    56.86961    89.89259    57.73486    90.40972
58.59835    90.93174    59.45669    91.46097    60.30645
91.99977    61.14425    92.55045    61.96666    93.11537
62.77030    93.69685    63.55175    94.29724    64.30761
94.91886    65.03447    95.56406    65.72892    96.23517
66.38757    96.93454    67.00700    97.66448    67.58382
98.47074    68.20060    99.29458    68.85865    100.1340
69.54589    100.9871    70.25024    101.8520    70.95962
102.7266    71.66196    103.6091    72.34519    104.4974
72.99722    105.3897    73.60597    106.2839    74.15938
107.1782    74.64535    108.0705    75.05183    108.9590
75.36672    109.8416    75.57796    110.7165    75.67345
111.5816    75.64114    112.2070    75.54066    112.8491
75.38528    113.5039    75.17875    114.1675    74.92480
114.8358    74.62715    115.5047    74.28954    116.1704
73.91571    116.8286    73.50939    117.4755    73.07431
118.1071    72.61420    118.7192    72.13281    119.3079
71.63385    119.8691    71.12107    120.3989    70.59820
120.8932    70.06897    121.3481    69.53711    121.7058
69.05543    122.0243    68.54285    122.3087    68.00318
122.5641    67.44024    122.7957    66.85785    123.0088
66.25982    123.2084    65.64996    123.3998    65.03210
123.5881    64.41005    123.7785    63.78763    123.9762
63.16865    124.1863    62.55693    124.4140    61.95629
124.6645    61.37054    124.9430    60.80350    125.2547
60.25898    125.8203    59.37940    126.4157    58.51649
127.0411    57.67027    127.6963    56.84073
```

APPENDIX III. MATRICES WITH NEAREST NEIGHBOUR DISTANCES, FOR ACROSS-GROUP COMPARISON AND FOR WITHIN-GROUP COMPARISON

	r1	r2	r3	r4	r5	r6	r7	r8	r9	r10	r11	r12	r13	r14	r15
b1	4.8157	3.4776	4.9345	3.7815	3.9872	3.7933	3.4895	4.1134	4.6666	4.5062	4.8661	4.7565	4.4160	4.1881	4.7041
b2	4.7529	3.5207	4.8875	3.7553	3.9618	3.8867	3.3978	4.1319	4.7308	4.5422	4.8513	4.7239	4.4595	4.2041	4.7540
b3	5.2269	3.9052	5.4000	4.2137	4.4477	4.2935	3.8845	4.6031	5.1372	4.9778	5.3089	5.1832	4.9228	4.6615	5.1973
b4	5.4258	4.2403	5.4181	4.4464	4.6343	4.6781	3.9692	4.8073	5.4698	5.2620	5.4753	5.4368	5.1256	4.9202	5.4716
b5	5.2701	3.9831	5.1700	4.1900	4.4116	4.2775	3.8558	4.5324	5.0712	4.9133	5.1720	5.1797	4.7702	4.5925	5.1453
b6	5.2102	3.8995	5.3026	4.1643	4.4005	4.2831	3.7904	4.5604	5.1299	4.9677	5.2644	5.1705	4.8589	4.6372	5.1607
b7	5.4269	4.1953	5.3141	4.3361	4.5583	4.5074	3.9527	4.7254	5.2992	5.1575	5.3198	5.3437	4.9915	4.7536	5.3916
b8	5.0078	3.8206	4.7853	3.9213	4.1084	4.0479	3.6306	4.2511	4.7474	4.6243	4.7911	4.8599	4.4394	4.2516	4.8625
b9	4.5830	3.4341	4.5580	3.5571	3.7436	3.7184	3.2382	3.8759	4.4987	4.3123	4.5561	4.5242	4.1556	3.9632	4.5344
b10	5.4436	4.2568	5.2431	4.3489	4.5288	4.5757	3.8594	4.6769	5.3613	5.1821	5.2655	5.3460	4.9516	4.7761	5.3957
b11	4.6774	3.5200	4.6902	3.6730	3.8713	3.8663	3.2317	4.0297	4.6892	4.4503	4.6676	4.6468	4.3626	4.1011	4.6900
b12	4.5933	3.4736	4.4147	3.5589	3.7440	3.8701	3.1394	3.8872	4.5520	4.3525	4.5391	4.5304	4.1399	3.9679	4.5610
b13	4.5132	3.2939	4.5880	3.5358	3.7300	3.6679	3.1976	3.8648	4.4833	4.3136	4.6230	4.4848	4.1934	3.9854	4.4887
b14	4.7983	3.6269	4.6455	3.8214	4.0083	3.9882	3.3832	4.1181	4.7372	4.5354	4.7725	4.7449	4.3949	4.1833	4.7573
b15	4.7639	3.5340	4.7865	3.7864	3.9856	3.9566	3.3808	4.1103	4.7438	4.5587	4.8628	4.7431	4.4101	4.1902	4.7371

A matrix with pairs of curves for “across group” comparison. The labels in the first column represent the 15 curves in the black set, and the labels in the first row represent the 15 curves in the red set (the two sets of curves are displayed in Figure 3-10). In total, 225 average nearest neighbour distance values are obtained. Each number represents the absolute difference between the two curves concerned, using the Nearest Neighbour technique (see Sections 3.3.6 and 3.3.7.1).

	r1	r2	r3	r4	r5	r6	r7	r8	r9	r10	r11	r12	r13	r14	r15
r1		1.2227	1.1722	0.8820	0.7977	1.5151	1.3611	0.6413	1.0976	0.7164	1.2863	0.5862	0.7278	0.7737	0.8350
r2			2.1959	0.7191	0.8719	0.7636	1.1860	0.9624	1.2431	1.1180	1.9125	1.2822	1.3374	0.9701	1.2405
r3				1.6056	1.3856	2.2129	1.5223	1.2690	1.5348	1.2556	0.6955	1.0644	1.0669	1.3065	1.4627
r4					0.3973	1.0427	0.8732	0.5305	1.1219	0.7470	1.3365	0.8728	0.8502	0.5054	1.0585
r5						1.2052	0.7966	0.4307	1.1931	0.7025	1.1474	0.7371	0.7020	0.5740	1.0356
r6							1.7702	1.2144	0.9333	1.1134	1.8979	1.3409	1.2133	1.2623	0.8510
r7								1.0311	1.8141	1.3939	1.4858	1.3854	1.3341	1.0712	1.6296
r8									1.0172	0.5349	1.0267	0.5082	0.5541	0.4331	0.8501
r9										0.6529	1.2113	0.8449	0.7192	0.8961	0.4674
r10											0.8897	0.4026	0.4122	0.5612	0.5783
r11												0.8608	0.9235	1.0615	1.2837
r12													0.3986	0.6375	0.6548
r13														0.6508	0.5613
r14															0.8431
r15															

A matrix with pairs of curves for “within group” comparison of the red curves (displayed in Figure 3-10). The labels both in the first row and in the first column represent 15 repetitions of the same stimulus, and the matrix illustrates which pairs of curves are compared. In total, 105 average nearest neighbour distance values are obtained. Each number represents the absolute difference between the two curves concerned, using the Nearest Neighbour technique (see Sections 3.3.6 and 3.3.7.2).

	b1	b2	b3	b4	b5	b6	b7	b8	b9	b10	b11	b12	b13	b14	b15
b1		0.7709	0.7407	1.1914	0.7269	0.7985	1.1762	0.8547	0.6674	1.1474	0.7946	0.8301	0.8566	0.7458	0.7390
b2			0.6409	0.9370	0.7070	0.5602	0.7813	0.7290	0.5091	0.9290	0.4665	0.7013	0.3189	0.5906	0.4564
b3				0.8109	0.7882	0.4695	0.9480	0.9882	0.9410	1.1001	0.9518	0.9934	0.7393	0.8798	0.7096
b4					0.6478	0.5486	0.6880	0.9683	1.1957	0.8140	1.1125	1.0136	0.9343	0.7878	0.7074
b5						0.5084	0.8163	0.6270	0.7722	0.7741	0.7408	0.6713	0.7488	0.4593	0.4511
b6							0.6305	0.7537	0.8487	0.7916	0.7681	0.8123	0.6263	0.5663	0.4288
b7								0.6709	0.8640	0.5686	0.8168	0.7770	0.8148	0.6489	0.6241
b8									0.6331	0.6426	0.6759	0.4587	0.8121	0.5955	0.5568
b9										0.9950	0.4006	0.5807	0.5510	0.5991	0.6092
b10											0.8819	0.6101	0.9843	0.6642	0.6620
b11												0.5637	0.5309	0.5315	0.5470
b12													0.7321	0.6032	0.5217
b13														0.6648	0.5030
b14															0.3831
b15															

A matrix with pairs of curves for “within group” comparison of the black curves (displayed in Figure 3-10). The labels both in the first row and in the first column represent 15 repetitions of the same stimulus, and the matrix illustrates which pairs of curves are compared. In total, 105 average nearest neighbour distance values are obtained. Each number represents the absolute difference between the two curves concerned, using the Nearest Neighbour technique (see Sections 3.3.6 and 3.3.7.2).

APPENDIX IV.

Essay on the Binomial Distribution

1. Definition and description

Binomial distribution is defined as “the distribution of outcomes from an experiment in which each of a number of independent trials results in one of two mutually exclusive outcomes” (Statistical Home Page. David C. Howell, Glossary). Binomial distribution describes the possible number of times that a particular event occurs in a sequence of observations. The distribution is specified by the number of observations (*n*) and the probability of occurrence (*p*). It is recorded how many times an event occurs in *n* repetitions of an experiment. For each repetition the event either occurs (and this is called a “success”) or it does not (this is called a “failure”).

Binomial distribution deals with discrete, and not continuous outcomes. If we plot this distribution, the *x* axis represents all the possible numbers of successes for *n* trials, and the *y* axis represents probabilities that these numbers of successes will occur. An example of a plot representing Binomial distribution is given in Figure 1. This distribution represents the situation with tossing a coin ten times. There are eleven outcomes, as all the cases are considered, including the one when the coin is not tossed at all. For all these trials (i.e., tossing the coin one time, or twice, or, for example, six times) the possibility of getting a head is calculated. We should mention that we take the head as a “success”, and then the tail comes out as “failure”. Rightly enough, in this case it does not matter which of the two outcomes to choose as a “success”, as there are only two of them, and the possibility of either of them occurring by pure chance is 50% (or 0.5). So here the “success”/”failure” labels are assigned arbitrarily. But in the case when there are, for example, three or four possible outcomes, it is important at which one of them we concentrate our attention, opposing it to all the others (this is relevant in our study).

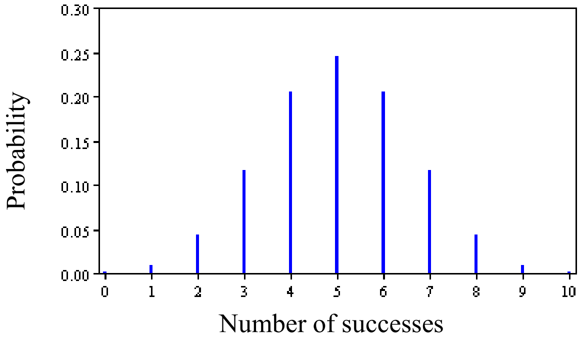


Figure 1. Binomial distribution when $n = 10$ and $p = 0.5$ (after *Online Statistics: A Multimedia Course of Studies*).

Binomial distribution can be used experimentally, to test hypotheses. It is possible to conduct a so-called binomial experiment when the following four conditions are satisfied: 1) there are *n* trials; 2) each trial results in a success or a failure; 3) the probability of a success, *p*, is constant from trial to trial; 4) the trials are independent. The outcome of this type of experiment is the number of successes, i.e., a count. The discrete variable *k* representing the number of successes is called a binomial random variable. The possible counts, *k* = 0, 1, 2, ..., *n*, and their associated probabilities define the binomial distribution, denoted by *f* (*n*, *p*). The mathematical formula for defining the binomial distribution is presented below:

$$\boxed{f(k;n,p) = \binom{n}{k} p^k (1-p)^{n-k}},$$

where $k = 0, 1, 2, \ldots, n$, and $\binom{n}{k} = \frac{n!}{k!(n-k)!}$

(after *Binomial Distribution. Wikipedia, The Free Encyclopedia*).

In this formula, *f* is the probability of *k* successes, *n* is the number of trials, and *p* is the probability of a success on any one trial. We should mention that while the *p* symbol occurs more or less standardly throughout the literature on binomial distribution, to denote the probability of a success on any one trial (the only exception I have found is the symbol *π* at *Online Statistics: A Multimedia Course of Studies*), the other constituents of the formula are represented by different symbols in different works. For example, *f* is also represented by *F* (e.g., Larson 1974), *B* (e.g., *Binomial Distribution. Department of Statistics, West Virginia University*), *b* (e.g., Hacking 2001; Mosteller *et al.* 1970), *P* (e.g., *Online Statistics: A Multimedia Course of Studies*) or *p* (e.g., Howell 2002); *N* sometimes stands for *n* (e.g., *Online Statistics: A Multimedia Course of Studies*; Weisstein*)*; *k* may be represented by *X* (e.g., Howell 2002), *x* (e.g., Mosteller *et al.* 1970), *n* (e.g., Weisstein), *r* (e.g., *Online Statistics: A Multimedia Course of Studies*) or *y* (e.g., Mendenhall 1987). Also, the symbol *q* is often used, to define the probability of a failure on any one trial, and it obviously equals $1 - p$.

Binomial distribution differs from many other distributions in that it is obtained mathematically, and not empirically. The *x* axis represents statistics (the number of successes as obtained in a given experiment) rather than individual observations or events. This distribution suits us in our purpose to find out whether trough occurrence in our subjects is different from random. We compare the actual, empirical number of occurrences with the mathematically defined distribution of chance occurrences, and we can then draw conclusions.

2. Statistics on quantitative distribution of tongue contour sequence patterns (used in Chapter 4 and Chapter 5)

In our case, the binomial random variable represents the number of trough occurrences, as opposed to all the other three tongue shape patterns taken together.

The aim is to compute probabilities (*f*) for a binomial random variable representing the number of troughs out of four different tongue shape patterns and visualise these quantities using a computer programme designed for this purpose. There are numerous resources on the web allowing for calculating this (see, e.g., *Binomial Distribution. Department of Statistics, West Virginia University*, or *StatBox 4.2*, in the reference list). I chose, rather arbitrarily, the Binomial Calculator available at *Online Statistics: A Multimedia Course of Studies* (http://psych.rice.edu/online_stat/java/binomialProb.html). The graphs produced by this Calculator will constitute all the following figures.

For each VCV sequence, there are 15 repetitions, or tokens (45 repetitions if we take all three subjects together)[6]. We assume that the four possible outcomes for tongue movement along a vertical measure bar (down-up = “trough”, up-down = “antitrough”, up-up = “continuous up”, down-down = “continuous down”) have equal chances of occurrence. We then expect equal chance probability of occurrence of all the four patterns, i.e., the probability of getting any of these patterns, including trough, is 0.25. In this description I only look at troughs, opposing them to all the other tongue shape patterns (non-troughs). However, the same procedure is easily applied to any of the other three patterns, with the purpose to see whether they occur by chance or not.

The conditions of the binomial experiment are assumed to be met: $n = 15$ (or 45) repetitions constitute the trials; each repetition results in one of two possible outcomes (trough or non-trough); the probability of trough occurrence p is 0.25 and is constant; the repetitions are independent.

2.1. Binomial Distribution, data for one subject and one VCV type

The number of troughs, k, is distributed as a binomial random variable with parameters $n = 15$ and $p = 0.25$. The distribution is presented in Figure 2.

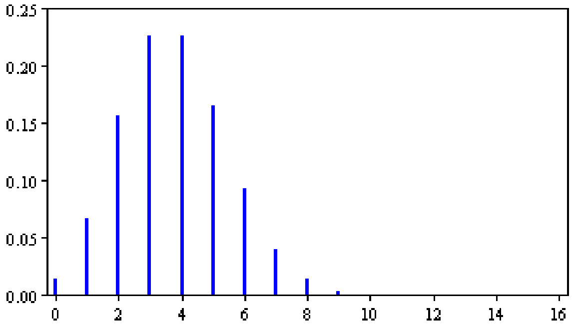


Figure 2. The graph presenting the probabilities of 16 possible outcomes, when $n = 15$ and $p = 0.25$.

The mean (or expected value) of a binomial distribution (μ) equals the number of trials n multiplied by the possibility of success for each trial p, i.e., $\mu = np$. So when $n = 15$ and $p = 0.25$, we can expect 3.75 (= 15 * 0.25) troughs if they occur randomly. Standard deviation of the binomial distribution is the square root of its variance, i.e., sqrt (npq). When $n = 15$ and $p = 0.25$ (q equals $1 - p$, as we mentioned earlier), the standard deviation is 1.68 = sqrt (15 * 0.25 * 0.75). If tongue patterns occur by chance, then approximately two-thirds of all the trials will fall within one standard deviation of the mean, i.e., there will be 3, 4, or 5 troughs.

---

[6] The calculations are described here only for these two cases: 15 tokens of a VCV, and 45 tokens of a VCV. In practice, there were a few more cases: 30 tokens (Chapter 4, voiced and voiceless consonants together, data from one subject), 90 tokens (Chapter 4, voiced and voiceless consonants together, data from three subjects), and 270 tokens (Chapter 4, all tokens, three subjects together). Mathematical probabilities for numbers of trough occurrences in these cases were calculated separately. Parameters used for the calculations are presented below, in Section 2.4.

The probabilities of the different trough occurrence numbers are given in Table 1. By plotting these numbers we are plotting the probability that a particular tongue contour sequence pattern (trough in our case) will occur by chance.

Number of Troughs, *k*	Probability (*f*) of *k* troughs occurring (to 4 decimal places)
0	0.0134
1	0.0668
2	0.1559
3	0.2252
4	0.2252
5	0.1651
6	0.0917
7	0.0393
8	0.0131
9	0.0034
10	0.0007
11	0.0001
12	0.0000
13	0.0000
14	0.0000
15	0.0000

Table 1. Mathematical probabilities (*f*) for *k* troughs out of 15 repetitions, assuming no bias.

We are interested in the probabilities of getting various numbers of troughs by chance. One possible case is when no troughs are produced. The probability of this event, calculated by the program, is 0.0134. It follows that there is a probability of 0.9866 (= 1 – 0.0134) of getting *at least* one trough. If we observe no troughs at all, there is a statistically significant probability that this situation is not accidental. The probability of occurrence of one trough in 15 repetitions is 0.0668, which is already above the commonly used threshold of significance of 0.05. When six troughs occur out of 15 repetitions, this is still at chance level. The probability of having seven troughs by chance is smaller than 0.05, so this is considered significant.

What is the lower limit of the number of troughs that we have to have, in order to be able to say that the probability of their occurrence by chance is less than 0.05? For that, we have to sum the possibilities of all the numbers starting from 7 (as 7 has the highest statistically significant *f* value, indicating that this number of troughs occurs not by chance). However, if we sum the probabilities of the numbers from 7 to 15, we will have the *f* value of 0.0566 as a result, which is more than 0.05. If we exclude the *f* for number 7 from this count, we will have a “decent” result of 0.0173, and this value will indicate that the answer to our question is that there should be eight or more troughs, in order to consider them occurring not at chance level.

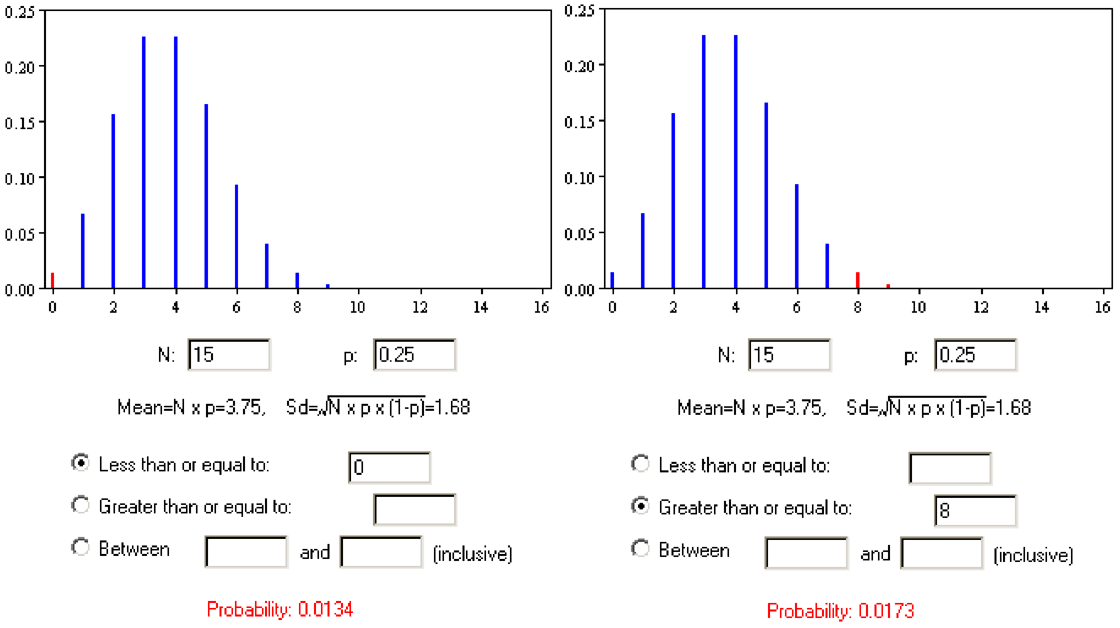
Probabilities of *k* or more trough occurrences are given in Table 2. There is a 100% probability that 0 or more troughs will occur; *f* values for one or more, two or more, and so on, up to seven or more troughs, are all greater than the widely accepted cut-off value of 0.05. Only starting from eight or more troughs we can say that the probability of chance occurrence of these numbers of troughs is low enough to accept that there is a statistically significant probability of their non-random occurrence.

Number of Troughs, k	Probability (f) of k or more troughs occurring (to 4 decimal places)
0	1
1	0.9866
2	0.9198
3	0.7639
4	0.5387
5	0.3135
6	0.1484
7	0.0566
8	0.0173
9	0.0042
10	0.0008
11	0.0001
12	0.0000
13	0.0000
14	0.0000
15	0.0000

Table 2. Mathematical probabilities (f) for k or more troughs out of 15 repetitions, assuming no bias.

When we compare Table 1 and Table 2, we see that all the f values in Table 2 are greater than the corresponding values in Table 1 (except for the last four rows, where f values are too small to see the difference in numbers rounded to four decimal places). In Table 2, f values are higher, and thus more "demanding", but they also give us more confidence in our results. For example, if we take the probability of exactly eight trough occurrences (Table 1), it is 0.0131, and the probability of eight or more trough occurrences (Table 2) is 0.0173. The former number shows us the probability of exactly eight troughs happening by chance, and this slightly exaggerates the significance of our results. The latter value, however, gives us the probability of chance occurrence of any of eight or more troughs, thus giving us the results we are really interested in.

To sum it up, when we have 15 repetitions ($n = 15$) and we assume chance occurrence of trough ($p = 0.25$), the situation when the number of troughs (k) has a probability of chance occurrence of less than 0.05 is the following: $k = 0$ and $k \geq 8$. This is graphically represented in Figure 3.



N: 15 p: 0.25

Mean=N x p=3.75, Sd=$\sqrt{N \times p \times (1-p)}$=1.68

● Less than or equal to: 0

○ Greater than or equal to:

○ Between and (inclusive)

Probability: 0.0134

N: 15 p: 0.25

Mean=N x p=3.75, Sd=$\sqrt{N \times p \times (1-p)}$=1.68

○ Less than or equal to:

● Greater than or equal to: 8

○ Between and (inclusive)

Probability: 0.0173

a) b)

Figure 3. Graphs representing the situation when the number of troughs (*k*) has a probability of less than 0.05 of being due to chance (when $n = 15$ and $p = 0.25$): a) $k = 0$; b) $k \geq 8$.

2.2. Binomial Distribution, data for three subjects taken together (one VCV type)

When the three subjects' productions are taken together, the theoretical distribution is presented in Figure 4.

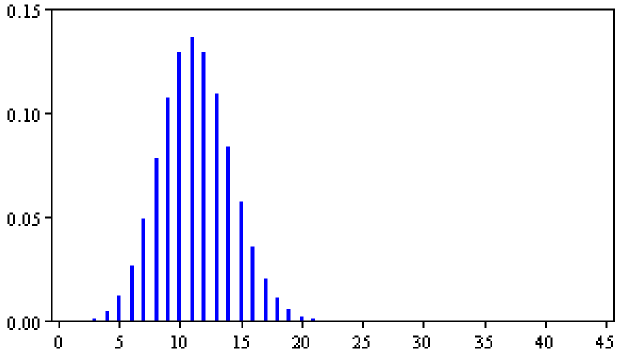
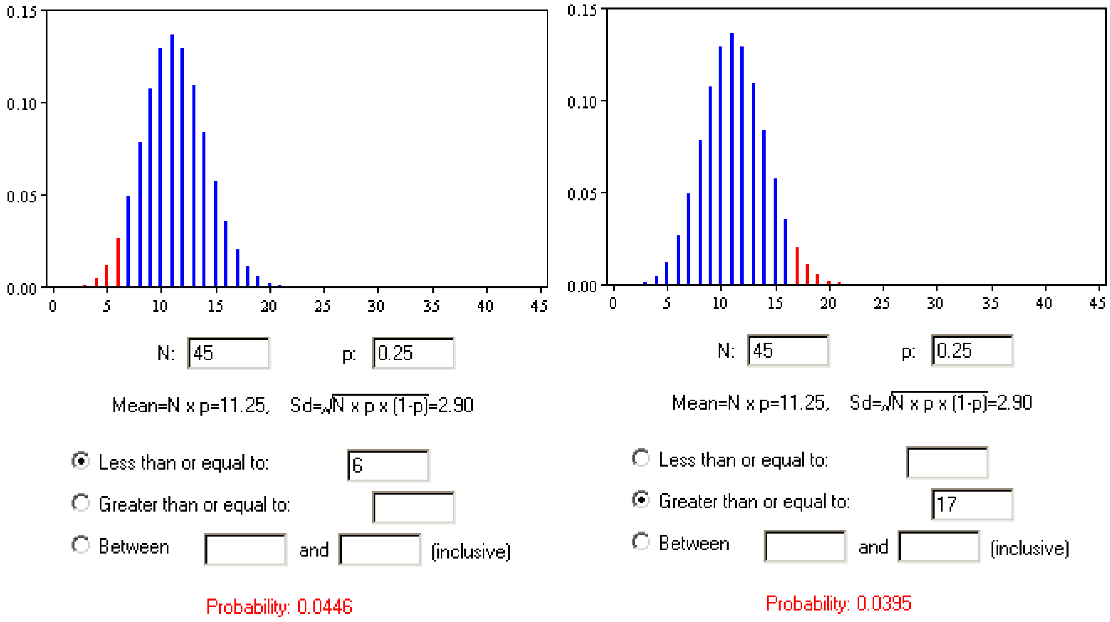


Figure 4. The graph presenting the probabilities of 46 possible outcomes, when $n = 45$ and $p = 0.25$.

In the case of 45 repetitions ($n = 45$), and with the same assumption of occurrence of trough $p = 0.25$, the numbers are the following. The situation when the number of troughs (*k*) have a low probability ($p < 0.05$) of being due to chance: $k \leq 6$ and $k \geq 17$. This is graphically represented in Figure 5.



N: 45 p: 0.25

Mean=N x p=11.25, Sd=√N x p x (1-p)=2.90

● Less than or equal to: 6
○ Greater than or equal to:
○ Between and (inclusive)

Probability: 0.0446

a)

N: 45 p: 0.25

Mean=N x p=11.25, Sd=√N x p x (1-p)=2.90

○ Less than or equal to:
● Greater than or equal to: 17
○ Between and (inclusive)

Probability: 0.0395

b)

Figure 5. Graphs representing the situation when the number of troughs (*k*) has a probability of less than 0.05 of being due to chance (when $n = 45$ and $p = 0.25$): a) $k \leq 6$; b) $k \geq 17$.

2.3. Bonferroni adjustment

In my work, in all the statistical calculation based on the binomial distribution, there are four theoretically possible tongue contour sequence patterns (trough, antitrough, continuous up and continuous down. So when I calculate mathematical probabilities for numbers of occurrence of these patterns in different conditions (i.e., different n values), I need to do a separate test for each of these four patterns, irrespective of the condition. When four separate tests are done, there is some chance fluctuation in the total experiment. As a result, the chance of finding at least one test statistically significant increases, as compared with doing just one test. Consequently, there are more chances to incorrectly declare a difference to be significant. In order to avoid this risk, the Bonferroni adjustment was used, to ensure that the overall risk for the four tests remains 0.05. The calculations were performed using an automatic procedure available for example on the *Simple Interactive Statistical Analysis* web page. The cut-off threshold for four separate tests was established at 0.0125.

2.4. Analysis parameters used in this work

Here, the *k* and *f* values are given, that were used in binomial experiments in this work.

In Tables 3 and 4, there are mathematical probabilities (*f*) for *k* or less troughs and *k* or more troughs out of 270 repetitions (used in Experiment 1, distribution of the four different patterns across speakers). In the binomial experiment with the Bonferroni adjustment applied, if $\boldsymbol{n = 270}$, the following *k* values have the probability of occurrence smaller than 0.0125:

$\boldsymbol{k \geq 85}$ (rate of occurrence significantly above chance);
$\boldsymbol{k \leq 51}$ (rate of occurrence significantly below chance).

Number of occurrences, *k*	Probability (*f*) of *k* or less patterns occurring (to 4 decimal places)
...40	**0.0000**
41	**0.0001**
42	**0.0001**
43	**0.0002**
44	**0.0004**
45	**0.0007**
46	**0.0011**
47	**0.0018**
48	**0.0029**
49	**0.0046**
50	**0.0071**
51	**0.0106**
52	0.0156
53	0.0224
54	0.0316
55...	0.0436

Table 3. Mathematical probabilities (*f*) for *k* or less patterns out of 270 repetitions, with $p = 0.25$, assuming no bias. The lines in bold script represent the numbers of occurrences being significantly below chance, and their probability values.

Number of occurrences, *k*	Probability (*f*) of *k* or more patterns occurring (to 4 decimal places)
…79	0.0628
80	0.0477
81	0.0357
82	0.0263
83	0.0191
84	0.0137
85	**0.0096**
86	**0.0067**
87	**0.0045**
88	**0.0031**
89	**0.0020**
90	**0.0013**
91	**0.0008**
92	**0.0005**
93	**0.0003**
94	**0.0002**
95	**0.0001**
96	**0.0001**
97…	**0.0000**

Table 4. Mathematical probabilities (*f*) for *k* or more patterns out of 270 repetitions, with $p = 0.25$, assuming no bias. The lines in bold script represent the numbers of occurrences being significantly above chance, and their probability values.

In Tables 5 and 6, there are mathematical probabilities (*f*) for *k* or less troughs and *k* or more troughs out of 90 repetitions (used in Experiment 1 and Experiment 2, distribution of the four different patterns across speakers, by vowel type). In the binomial experiment with the Bonferroni adjustment applied, if ***n* = 90**, the following *k* values have the probability of occurrence smaller than 0.0125:
***k* ≥ 33** (rate of occurrence significantly above chance);
***k* ≤ 13** (rate of occurrence significantly below chance).

Number of occurrences, *k*	Probability (*f*) of *k* or less patterns occurring (to 4 decimal places)
...7	**0.0000**
8	**0.0001**
9	**0.0003**
10	**0.0008**
11	**0.0022**
12	**0.0051**
13	**0.0110**
14	0.0218
15...	0.0399

Table 5. Mathematical probabilities (*f*) for *k* or less patterns out of 90 repetitions, with $p = 0.25$, assuming no bias. The lines in bold script represent the numbers of occurrences being significantly below chance, and their probability values.

Number of occurrences, *k*	Probability (*f*) of *k* or more patterns occurring (to 4 decimal places)
…29	0.0748
30	0.0473
31	0.0287
32	0.0167
33	**0.0093**
34	**0.0049**
35	**0.0025**
36	**0.0012**
37	**0.0006**
38	**0.0003**
39	**0.0001**
40…	**0.0000**

Table 6. Mathematical probabilities (*f*) for *k* or more patterns out of 90 repetitions, with $p = 0.25$, assuming no bias. The lines in bold script represent the numbers of occurrences being significantly above chance, and their probability values.

In Tables 7 and 8, there are mathematical probabilities (*f*) for *k* or less troughs and *k* or more troughs out of 45 repetitions (used in Experiment 1 and Experiment 2, distribution of the four different patterns across speakers, by VCV type). In the

binomial experiment with the Bonferroni adjustment applied, if ***n* = 45**, the following *k* values have the probability of occurrence smaller than 0.0125:
***k* ≥ 19** (rate of occurrence significantly above chance);
***k* ≤ 4** (rate of occurrence significantly below chance).

Number of occurrences, *k*	Probability (*f*) of *k* or less patterns occurring (to 4 decimal places)
0	**0.0000**
1	**0.0000**
2	**0.0003**
3	**0.0016**
4	**0.0059**
5	0.0179
6…	0.0446

Table 7. Mathematical probabilities (*f*) for *k* or less patterns out of 45 repetitions, with $p = 0.25$, assuming no bias. The lines in bold script represent the numbers of occurrences being significantly below chance, and their probability values.

Number of occurrences, *k*	Probability (*f*) of *k* or more patterns occurring (to 4 decimal places)
...16	0.0753
17	0.0395
18	0.0191
19	**0.0085**
20	**0.0035**
21	**0.0013**
22	**0.0005**
23	**0.0001**
24…	**0.0000**

Table 8. Mathematical probabilities (*f*) for *k* or more patterns out of 45 repetitions, with $p = 0.25$, assuming no bias. The lines in bold script represent the numbers of occurrences being significantly above chance, and their probability values.

In Tables 9 and 10, there are mathematical probabilities (*f*) for *k* or less troughs and *k* or more troughs out of 30 repetitions (used in Experiment 1, distribution of the four different patterns by vowel type and by subject). In the binomial experiment with the Bonferroni adjustment applied, if ***n* = 30**, the following *k* values have the probability of occurrence smaller than 0.0125:
***k* ≥ 14** (rate of occurrence significantly above chance);
***k* ≤ 2** (rate of occurrence significantly below chance).

Number of occurrences, *k*	Probability (*f*) of *k* or less patterns occurring (to 4 decimal places)
0	**0.0002**
1	**0.0020**
2	**0.0106**
3…	0.0374

Table 9. Mathematical probabilities (*f*) for *k* or less patterns out of 30 repetitions, with $p = 0.25$, assuming no bias. The lines in bold script represent the numbers of occurrences being significantly below chance, and their probability values.

Number of occurrences, *k*	Probability (*f*) of *k* or more patterns occurring (to 4 decimal places)
…11	0.1057
12	0.0507
13	0.0216
14	**0.0082**
15	**0.0027**
16	**0.0008**
17	**0.0002**
18	**0.0001**
19…	**0.0000**

Table 10. Mathematical probabilities (*f*) for *k* or more patterns out of 30 repetitions, with $p = 0.25$, assuming no bias. The lines in bold script represent the numbers of occurrences being significantly above chance, and their probability values.

In Tables 11 and 12, there are mathematical probabilities (*f*) for *k* or less troughs and *k* or more troughs out of 15 repetitions (used in Experiment 1 and Experiment 2, distribution of the four different patterns by VCV type and by subject). In the binomial experiment with the Bonferroni adjustment applied, if $\boldsymbol{n = 15}$, the following *k* values have the probability of occurrence smaller than 0.0125:
$\boldsymbol{k \geq 9}$ (rate of occurrence significantly above chance).

Number of occurrences, *k*	Probability (*f*) of *k* or less patterns occurring (to 4 decimal places)
0	0.0134
1…	0.0802

Table 11. Mathematical probabilities (*f*) for *k* or less patterns out of 15 repetitions, with $p = 0.25$, assuming no bias. Note that there are no lines in bold script; this means that no numbers of occurrences are significantly below chance.

Number of occurrences, *k*	Probability (*f*) of *k* or more patterns occurring (to 4 decimal places)
...6	0.1484
7	0.0566
8	0.0173
9	**0.0042**
10	**0.0008**
11	**0.0001**
12...	**0.0000**

Table 12. Mathematical probabilities (*f*) for *k* or more patterns out of 15 repetitions, with $p = 0.25$, assuming no bias. The lines in bold script represent the numbers of occurrences being significantly above chance, and their probability values.

In Tables 13 and 14, there are mathematical probabilities (*f*) for *k* or less troughs and *k* or more troughs out of 100 repetitions (used in Experiment 2 for analysing the data presented in Vazquez Alvarez et al., 2004; distribution of the four different patterns across speakers, by VCV type). In the binomial experiment with the Bonferroni adjustment applied, if $\boldsymbol{n} = \mathbf{100}$, the following *k* values have the probability of occurrence smaller than 0.0125:

$\boldsymbol{k} \geq$ **36** (rate of occurrence significantly above chance);

$\boldsymbol{k} \leq$ **15** (rate of occurrence significantly below chance).

Number of occurrences, *k*	Probability (*f*) of *k* or less patterns occurring (to 4 decimal places)
...9	**0.0000**
10	**0.0001**
11	**0.0004**
12	**0.0010**
13	**0.0025**
14	**0.0054**
15	**0.0111**
16	0.0211
17	0.0376
18...	0.0630

Table 13. Mathematical probabilities (*f*) for *k* or less patterns out of 100 repetitions, with $p = 0.25$, assuming no bias. The lines in bold script represent the numbers of occurrences being significantly below chance, and their probability values.

Number of occurrences, *k*	Probability (*f*) of *k* or more patterns occurring (to 4 decimal places)
...32	0.0693
33	0.0446
34	0.0276
35	0.0164
36	**0.0094**
37	**0.0052**
38	**0.0027**
39	**0.0014**
40	**0.0007**
41	**0.0003**
42	**0.0001**
43	**0.0001**
44...	**0.0000**

Table 14. Mathematical probabilities (*f*) for *k* or more patterns out of 100 repetitions, with $p = 0.25$, assuming no bias. The lines in bold script represent the numbers of occurrences being significantly above chance, and their probability values.

In Tables 15 and 16, there are mathematical probabilities (*f*) for *k* or less troughs and *k* or more troughs out of 50 repetitions (used in Experiment 2 for analysing the data presented in Vazquez Alvarez et al., 2004; distribution of the four different patterns across speakers, by vowel type). In the binomial experiment with the Bonferroni adjustment applied, if $\boldsymbol{n} = \mathbf{50}$, the following *k* values have the probability of occurrence smaller than 0.0125:
$\boldsymbol{k} \geq$ **21** (rate of occurrence significantly above chance);
$\boldsymbol{k} \leq$ **5** (rate of occurrence significantly below chance).

Number of occurrences, *k*	Probability (*f*) of *k* or less patterns occurring (to 4 decimal places)
0	**0.0000**
1	**0.0000**
2	**0.0001**
3	**0.0005**
4	**0.0021**
5	**0.0070**
6	0.0194
7	0.0453
8...	0.0916

Table 15. Mathematical probabilities (*f*) for *k* or less patterns out of 50 repetitions, with $p = 0.25$, assuming no bias. The lines in bold script represent the numbers of occurrences being significantly below chance, and their probability values.

Number of occurrences, *k*	Probability (*f*) of *k* or more patterns occurring (to 4 decimal places)
...17	0.0983
18	0.0551
19	0.0287
20	0.0139
21	**0.0063**
22	**0.0026**
23	**0.0010**
24	**0.0004**
25	**0.0001**
26...	**0.0000**

Table 16. Mathematical probabilities (*f*) for *k* or more patterns out of 50 repetitions, with $p = 0.25$, assuming no bias. The lines in bold script represent the numbers of occurrences being significantly above chance, and their probability values.

2.5. An example from the data

Here, an illustration is presented of applying the method described above to some data from Experiment 2 (Chapter 5). For example, when we look at the results from all the subjects together (Figure 5-8), there occur 43 troughs in /uhu/ sequences, 41 troughs in /ihi/ sequences, and 11 troughs in /aha/ sequences. Looking at the previous section, we can say that the first two numbers fall within the range of *k* values for the rate of trough occurrence being significantly above chance (i.e., $k \leq 4$ and $k \geq 19$), $p < 0.001$. Thus the rate of trough occurrence in the two high vowel contexts demonstrates a significant discontinuity in coarticulation occurring in the two high vowels. In the open vowel /a/ context, the rate of occurrence of troughs is at chance level. None of the patterns in /aha/ sequences has the rate of occurrence above chance level. This means that there is no significant discontinuity in coarticulation in /aha/ sequences.

References

Berrie's Statistics Page. http://www.berrie.dds.nl/binomial.html

Binomial Calculator. http://psych.rice.edu/online_stat/java/binomialProb.html

Binomial Distribution. Department of Statistics, West Virginia University. http://www.stat.wvu.edu/SRS/Modules/Binomial/binomial.html

Binomial Distribution. Wikipedia, The Free Encyclopedia. http://en.wikipedia.org/wiki/Binomial_distribution

George, D. & Mallery, P. (2003). *SPSS for Windows Step by Step: A Simple Guide and Reference. 11.0 Update*. 4th edn. Boston, New York, San Francisco: Pearson Edication.

Hacking, I. (2001). *An Introduction to Probability and Inductive Logic*. Cambridge: Cambridge University Press.

Howell, D. C. (2002). *Statistical Methods for Psychology*. 5th edn. Pacific Grove, CA: Duxbury.

Larson, H. J. (1974). *Introduction to Probability Theory and Statistical Inference*. 2nd edn. New York, London, Sydney, Toronto: Jogh Wiley & Sons.

Mendenhall, W. (1987). *Introduction to Probability and Statistics*. Boston: Duxbury.

Mosteller, F., Rourke, R. E. K. & Thomas, G. B. (1970). *Probability With Statistical Applications*. 2nd edn. London: Addison-Wesley Publishing Company.

Online Statistics: A Multimedia Course of Studies. David Lane, Departments of Psychology and Statistics, Rice University. http://psych.rice.edu/online_stat/

Simple Interactive Statistical Analysis. http://home.clara.net/sisa/bonhlp.htm

StatBox 4.2: A Statistics Toolbox for MATLAB.
http://www.statsci.org/matlab/statbox.html

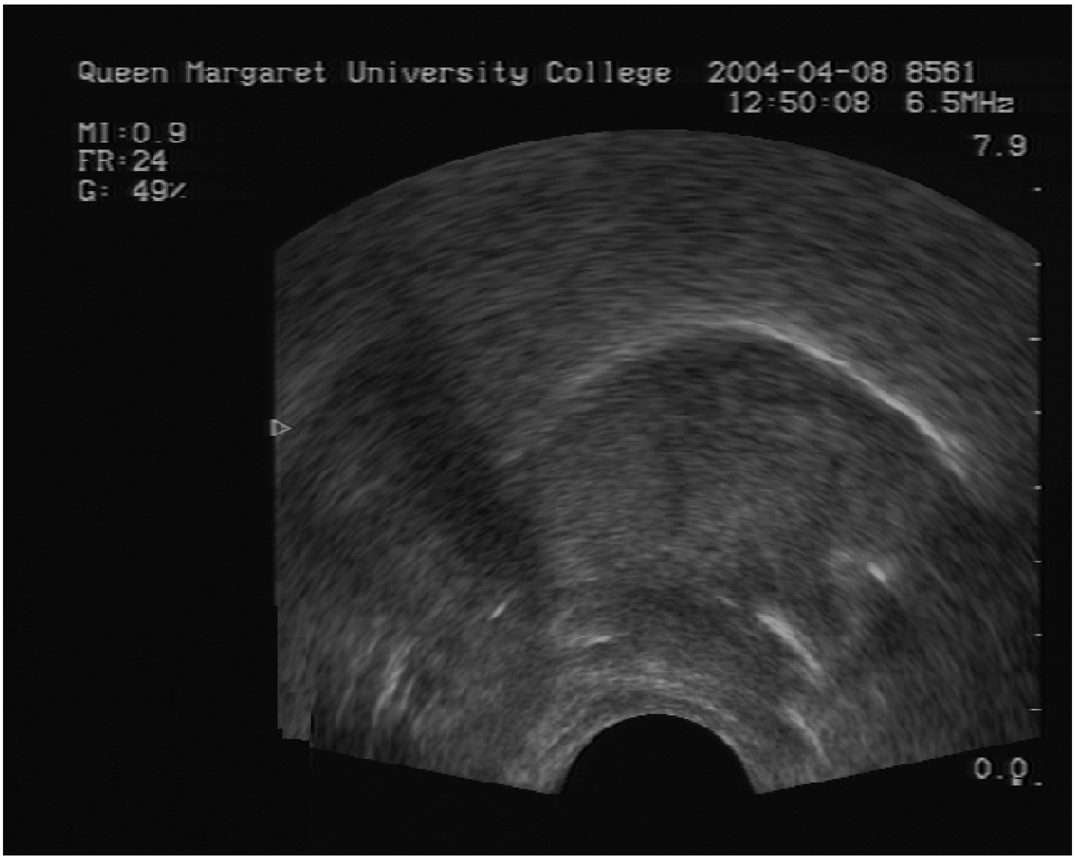
Statistical Home Page. David C. Howell.
http://www.uvm.edu/~dhowell/StatPages/StatHomePage.html

Weisstein, E. W. "Binomial Distribution." From *MathWorld*--A Wolfram Web Resource. http://mathworld.wolfram.com/BinomialDistribution.html

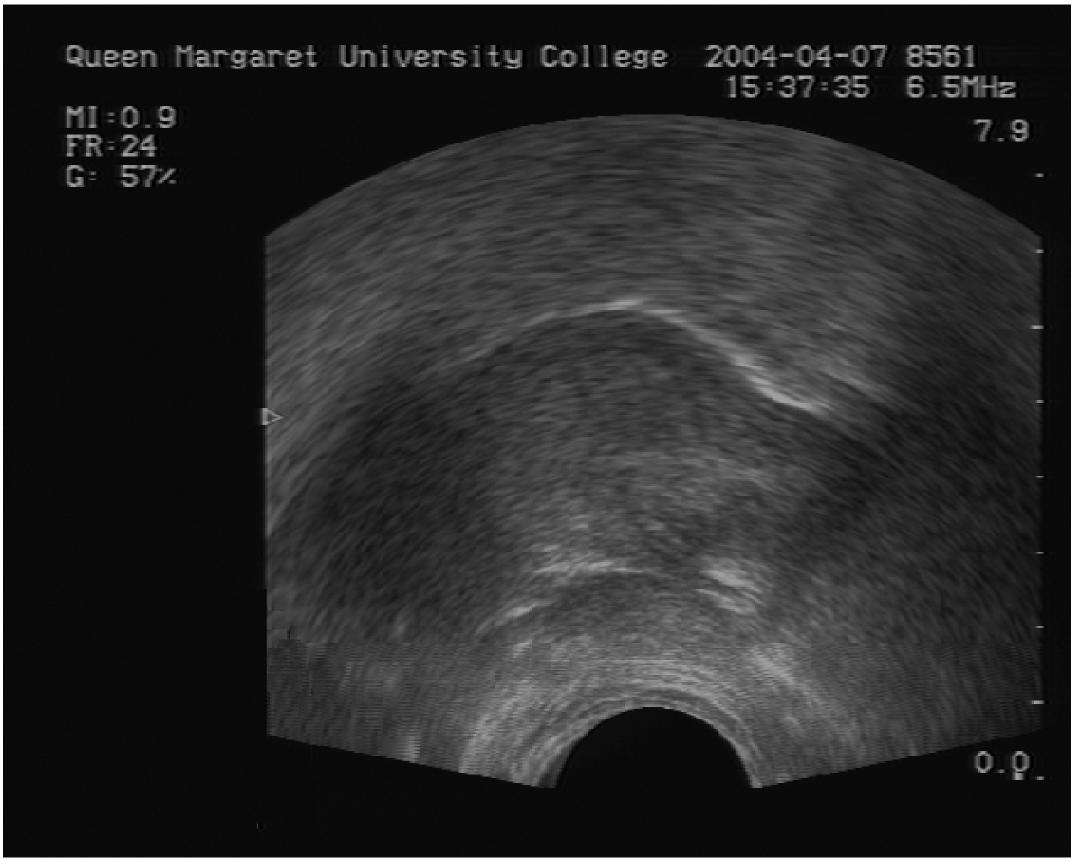
APPENDIX V. MIDSAGITTAL ULTRASOUND SCANS OF THE TONGUE AT SPEECH REST POSITION, FOR EXPERIMENT 1 AND EXPERIMENT 2

APPENDIX V-1. Experiment 1

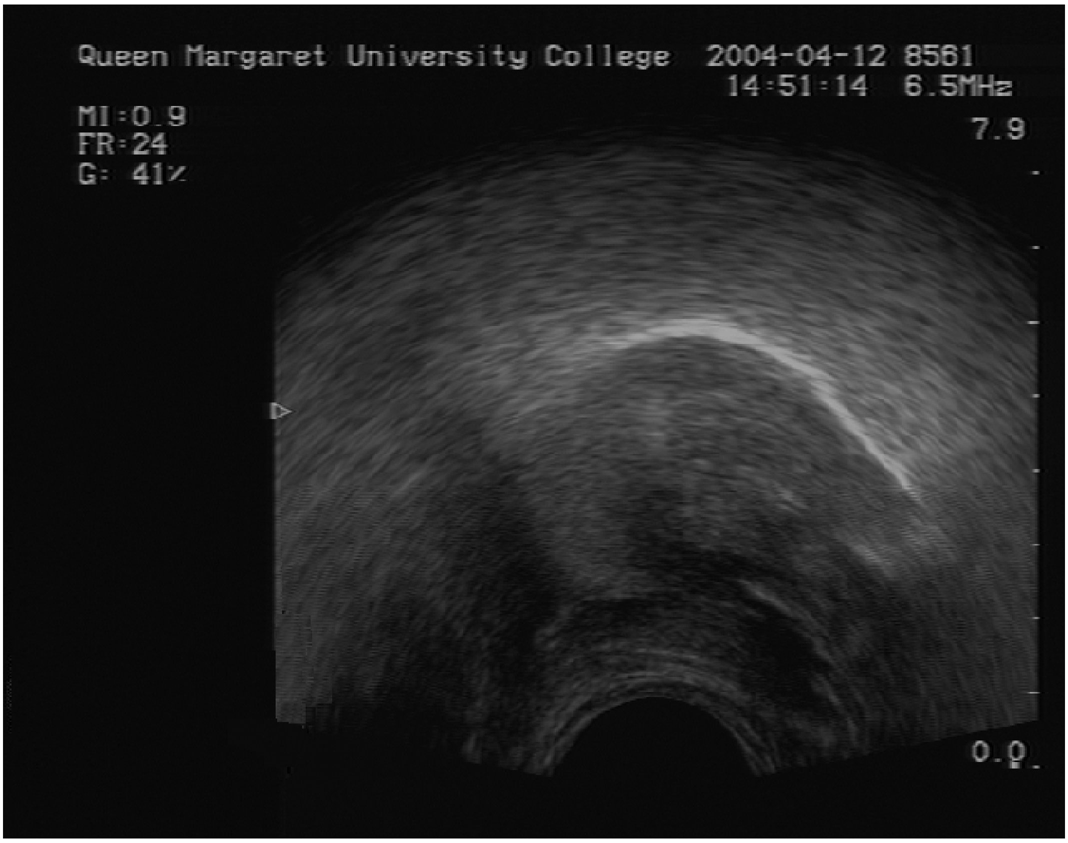
Subject S1. The tongue tip is on the right.



A midsagittal ultrasound scan of the tongue at speech rest position. Subject S2, Experiment 1. The tongue tip is on the right.

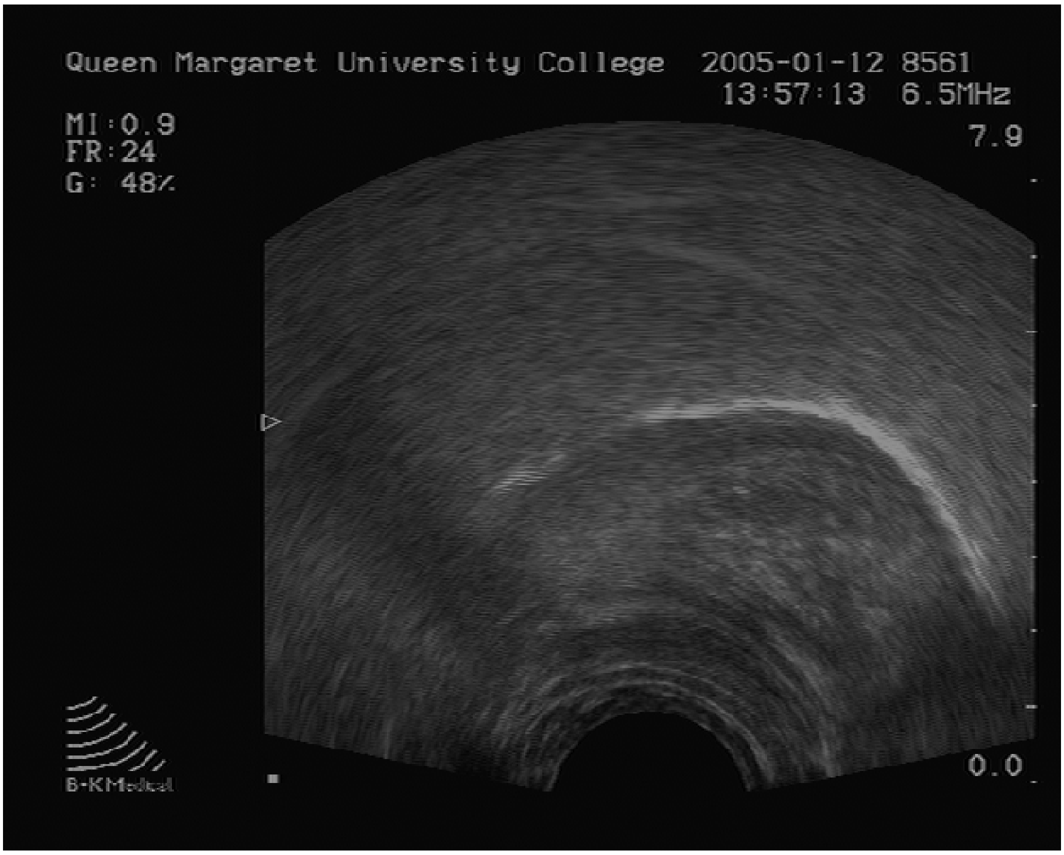


A midsagittal ultrasound scan of the tongue at speech rest position. Subject S3, Experiment 1. The tongue tip is on the right.

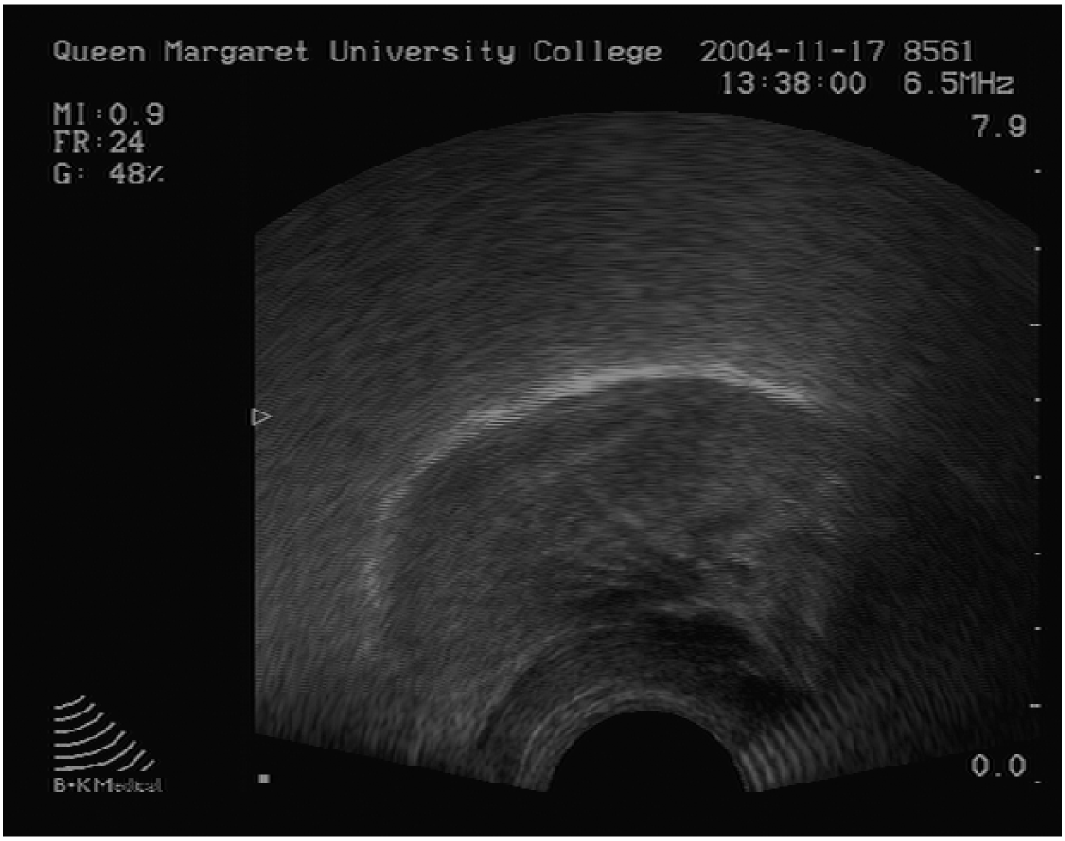


APPENDIX V-2. Experiment 2

A midsagittal ultrasound scan of the tongue at speech rest position. Subject S1, Experiment 2. The tongue tip is on the right.



A midsagittal ultrasound scan of the tongue at speech rest position. Subject S2, Experiment 2. The tongue tip is on the right.



A midsagittal ultrasound scan of the tongue at speech rest position. Subject S3, Experiment 2. The tongue tip is on the right.

APPENDIX VI. MATRICES WITH ACROSS-GROUP NEAREST NEIGHBOUR DISTANCES, FOR CALCULATING THE COARTICULATION RESISTANCE COEFFICIENT

APPENDIX VI-1.

Distances between t_i and t_a curves

A matrix with distances for calculating the Coarticulation Resistance Coefficient (CRC) of the consonant /t/ for subject S2 (cf. Section 7.3.1). The labels in the first column represent the 15 t_i curves, and the labels in the first row represent the 15 t_a curves (the two sets of curves are displayed in Figure 7-2). In total, 225 average nearest neighbour distance values are obtained. Each number represents the absolute difference between the two curves concerned, using the Nearest Neighbour technique (see Sections 3.4.6 and 3.4.7.1).

	a1	a2	a3	a4	a5	a6	a7	a8	a9	a10	a11	a12	a13	a14	a15
i1	6.0898	5.4049	5.8817	5.4338	6.4313	4.6880	4.9191	5.7324	5.9773	5.8546	5.8869	5.9965	5.7832	6.2081	5.2893
i2	6.1585	5.4007	5.9107	5.4278	6.4718	4.6675	4.8679	5.7672	6.1315	5.8175	5.8717	5.9943	5.8460	6.2805	5.3367
i3	5.0361	4.3226	4.7517	4.2980	5.3835	3.6139	3.8242	4.7656	4.8917	4.8127	4.8428	4.9734	4.7126	5.1339	4.2507
i4	5.6104	4.9620	5.3198	4.9358	5.9390	4.1406	4.4074	5.4110	5.5205	5.3826	5.5555	5.5477	5.2943	5.7530	4.9243
i5	5.2134	4.4183	4.8633	4.4409	5.4861	3.7127	3.9298	4.8800	5.0791	4.8814	4.9992	5.0381	4.8518	5.2321	4.4311
i6	5.6127	4.9300	5.3965	4.9329	5.9417	4.2067	4.4239	5.3208	5.5130	5.3895	5.4049	5.5234	5.3349	5.7428	4.7670
i7	5.6313	4.9492	5.3916	4.8968	5.9632	4.1860	4.3835	5.3720	5.5399	5.4035	5.4145	5.5517	5.3461	5.7791	4.8539
i8	4.0806	3.3297	3.5632	3.3106	4.3358	2.5903	2.9182	3.9589	3.9126	3.6976	4.0500	3.8636	3.6167	3.9353	3.4754
i9	4.4661	3.7364	4.1083	3.6733	4.7675	3.0135	3.2301	4.2672	4.3252	4.2058	4.3588	4.3782	4.0818	4.4880	3.7906
i10	4.6844	3.9893	4.3995	3.9521	4.9616	3.2451	3.4686	4.4037	4.4873	4.4050	4.5538	4.5529	4.3565	4.7373	3.9869
i11	4.6481	3.9499	4.3429	3.8843	4.8923	3.1391	3.3974	4.3654	4.4375	4.3781	4.4921	4.4730	4.3483	4.6989	3.9048
i12	4.3855	3.5948	3.8679	3.6169	4.6643	2.8995	3.1864	4.2354	4.2511	4.0271	4.3357	4.2234	3.8980	4.2753	3.7809
i13	4.7909	4.0044	4.4255	4.0458	5.1067	3.3049	3.5496	4.5585	4.6904	4.4411	4.6554	4.5894	4.4268	4.7818	4.0897
i14	5.3363	4.6893	5.0867	4.6663	5.7231	3.9617	4.2012	5.1278	5.1852	5.1461	5.2065	5.2935	5.0664	5.4610	4.5921
i15	4.7839	4.1293	4.4162	4.0528	5.0664	3.3631	3.6417	4.6568	4.5948	4.5831	4.7218	4.7211	4.4379	4.8068	4.0508

APPENDIX VI-2. Distances between h_i and h_a curves

	a1	a2	a3	a4	a5	a6	a7	a8	a9	a10	a11	a12	a13	a14	a15
i1	7.7708	7.8259	7.9389	7.6371	7.7088	7.4878	7.8026	7.7217	7.9074	8.1191	7.9501	7.7020	7.8721	8.0233	7.8851
i2	7.6402	7.7127	7.8742	7.4897	7.6378	7.4653	7.7263	7.5877	7.7752	7.9632	7.7937	7.5993	7.7644	7.9090	7.6976
i3	8.3861	8.4535	8.6602	8.2442	8.4229	8.2374	8.5712	8.3366	8.5074	8.7041	8.5403	8.3637	8.5118	8.6503	8.4538
i4	7.9205	8.0115	8.2090	7.7756	7.9519	7.7867	8.0576	7.8660	8.0460	8.2549	8.0879	7.8975	8.0378	8.2082	7.9607
i5	8.4043	8.5187	8.6008	8.2071	8.4071	8.2478	8.4651	8.3976	8.5138	8.6846	8.5539	8.3855	8.5543	8.7128	8.4129
i6	8.4467	8.4378	8.4976	8.1830	8.3287	8.1470	8.4690	8.3630	8.5394	8.7298	8.5582	8.3977	8.5365	8.6885	8.5229
i7	8.3805	8.4859	8.5863	8.2626	8.3915	8.1617	8.4881	8.3593	8.5176	8.6975	8.5709	8.3485	8.5298	8.6578	8.4678
i8	8.4946	8.5583	8.6904	8.3591	8.4694	8.2569	8.5967	8.4446	8.6214	8.7881	8.6436	8.4382	8.5989	8.7394	8.6001
i9	7.9380	8.0210	8.1473	7.7986	7.9115	7.7211	8.0438	7.9167	8.0830	8.2558	8.0973	7.9017	8.0701	8.1946	8.0360
i10	8.0000	8.1092	8.2223	7.8984	8.0105	7.8317	8.1284	7.9663	8.1682	8.2994	8.1665	7.9974	8.1927	8.2597	8.0668
i11	8.8164	8.8533	8.9727	8.6665	8.7657	8.5357	8.9062	8.7937	8.9501	9.1211	8.9888	8.7795	8.9597	9.0566	8.9198
i12	8.0349	8.1182	8.1616	7.8504	7.9704	7.7743	8.0608	8.0096	8.1629	8.3406	8.1942	7.9812	8.1552	8.3178	8.1285
i13	8.3134	8.4260	8.5204	8.1782	8.3231	8.1290	8.4275	8.3051	8.4553	8.6163	8.4692	8.2988	8.4581	8.5702	8.3818
i14	8.7687	8.8562	8.9771	8.6360	8.7819	8.5795	8.9231	8.7223	8.8903	9.0684	8.9075	8.7288	8.8781	9.0094	8.8571
i15	8.4916	8.5761	8.5591	8.2674	8.4042	8.2252	8.4737	8.4661	8.6145	8.7877	8.6466	8.4561	8.6379	8.8104	8.5392

A matrix with distances for calculating the Coarticulation Resistance Coefficient (CRC) of the consonant /h/ for subject S2 (cf. Section 7.3.1). The labels in the first column represent the 15 h_i curves, and the labels in the first row represent the 15 h_a curves. In total, 225 average nearest neighbour distance values are obtained. Each number represents the absolute difference between the two curves concerned, using the Nearest Neighbour technique (see Sections 3.4.6 and 3.4.7.1).

APPENDIX VI-3. Distances between $a1_t$ and $a1_k$ curves

	k1	k2	k3	k4	k5	k6	k7	k8	k9	k10	k11	k12	k13	k14	k15
t1	1.2507	1.6190	1.6025	1.0960	1.8683	1.5870	1.3573	1.1483	1.1515	1.3131	1.7394	1.3099	1.4632	1.4962	1.4707
t2	0.9268	1.1922	1.2136	0.6975	1.4828	1.2755	1.0897	0.9773	0.8325	0.9990	1.3050	1.0340	1.0995	1.1139	0.9388
t3	0.9927	0.9406	0.9681	0.7207	1.2435	1.2140	1.1102	1.2374	0.9008	1.2896	1.1123	1.1726	0.9711	1.1003	0.8715
t4	0.9744	1.0554	1.1297	0.6947	1.3824	1.2359	1.1251	1.1186	0.8869	1.1814	1.2195	1.1593	1.1164	1.0458	0.7285
t5	1.1099	1.3138	1.4214	0.9700	1.7751	1.4805	1.3195	1.2658	1.0339	1.2810	1.5516	1.5008	1.2855	1.2922	1.1640
t6	1.1408	1.5158	1.4670	0.9988	1.7127	1.4281	1.2294	0.9801	1.0154	1.1475	1.5933	1.1735	1.3814	1.3899	1.2912
t7	0.6670	1.0885	0.9653	0.5574	1.3525	0.9638	0.8255	0.7450	0.6726	0.9851	1.1694	1.0853	0.9670	0.9051	0.9227
t8	1.1300	1.3501	1.3287	1.1244	1.6644	1.4373	1.4929	1.0704	0.9046	1.5518	1.3560	1.6475	1.1722	1.1801	0.9658
t9	1.0943	1.4054	1.2937	0.8292	1.6116	1.3382	1.1874	0.8429	0.9828	1.3191	1.4136	1.2244	1.2284	1.2443	1.0409
t10	1.0981	1.4143	1.3123	0.9457	1.6317	1.4613	1.3618	1.0316	0.9538	1.4379	1.4513	1.3823	1.2150	1.2474	0.9444
t11	1.2239	1.4848	1.4877	0.9478	1.7279	1.5365	1.3447	1.1799	1.1086	1.3089	1.5896	1.2580	1.4260	1.3601	1.0977
t12	0.7880	1.1831	1.2122	0.8062	1.5577	1.1784	1.1893	0.9979	0.6626	1.0753	1.3197	1.3461	1.0617	0.9819	0.8919
t13	0.9110	1.0545	1.0816	0.6498	1.3536	1.1981	1.0928	1.1348	0.8360	1.1391	1.1789	1.1493	1.0799	1.0062	0.7240
t14	0.8133	1.1096	1.1011	0.6477	1.4153	1.1375	1.0263	0.9783	0.7689	0.9881	1.2109	1.0564	1.0131	0.9803	0.7911
t15	1.0403	1.2959	1.3709	0.8592	1.6488	1.3852	1.2753	1.0948	0.8796	1.1516	1.3999	1.2754	1.2305	1.1453	0.8700

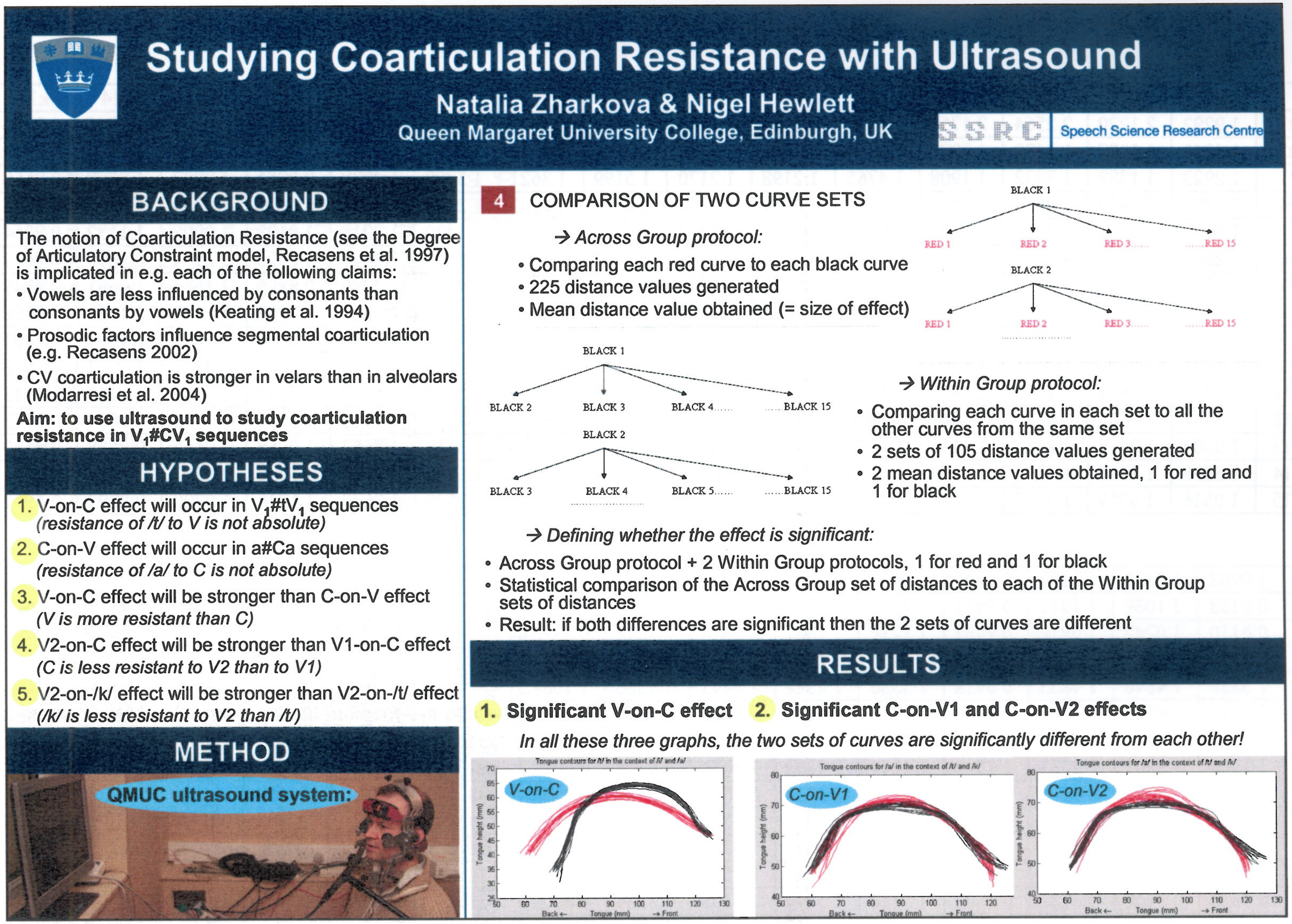
A matrix with distances for calculating the Coarticulation Resistance Coefficient (CRC) of V1 for subject S2 (cf. Section 7.3.2). The labels in the first column represent the 15 $a1_t$ curves, and the labels in the first row represent the 15 $a1_k$ curves. In total, 225 average nearest neighbour distance values are obtained. Each number represents the absolute difference between the two curves concerned, using the Nearest Neighbour technique (see Sections 3.4.6 and 3.4.7.1).

APPENDIX VI-4. Distances between $a2_t$ and $a2_k$ curves

	k1	k2	k3	k4	k5	k6	k7	k8	k9	k10	k11	k12	k13	k14	k15
t1	1.9292	2.1220	2.2564	2.1750	2.3224	2.1985	1.7626	2.2092	1.8413	2.0392	2.1637	1.9478	2.0071	1.8845	1.8725
t2	1.4340	1.1987	1.1482	1.3275	1.5854	1.3285	1.3171	1.5596	1.4766	1.7240	1.5535	1.0357	1.4034	1.1685	1.1012
t3	1.2923	1.1382	1.3032	1.1908	1.4765	1.2198	1.2170	1.5199	1.3621	1.5389	1.4232	0.9851	1.2807	1.0698	0.9568
t4	1.0701	0.9889	1.0533	1.0802	1.1434	1.0901	1.2850	1.4385	1.0202	1.2916	1.3031	1.0471	1.0291	0.8028	0.7506
t5	1.0370	1.0698	1.5080	1.3125	1.2721	1.1830	0.8493	1.3142	1.0282	1.1420	1.2213	1.0128	1.0908	1.1132	0.8254
t6	1.2967	1.1873	1.2679	1.1952	1.3670	1.2668	1.3408	1.5767	1.2670	1.5569	1.4447	1.0925	1.2784	0.9605	0.8617
t7	1.1960	0.9495	1.2272	1.1247	1.3748	1.0810	1.0778	1.2966	1.2821	1.4215	1.2999	0.8825	1.1667	1.0661	0.8849
t8	1.1690	1.3502	1.4970	1.4633	1.3686	1.4514	1.3488	1.6369	1.0327	1.3727	1.5037	1.3753	1.2682	1.1262	1.0198
t9	1.4781	1.5348	1.6610	1.6336	1.6958	1.6249	1.3835	1.7525	1.4534	1.7423	1.6885	1.3459	1.5357	1.4001	1.2491
t10	1.4721	1.4968	1.7521	1.6499	1.6663	1.6162	1.3471	1.7237	1.4192	1.6610	1.6382	1.3744	1.4772	1.3922	1.2033
t11	1.3012	1.4614	1.7092	1.4659	1.5935	1.5458	1.3149	1.6917	1.3059	1.5429	1.5189	1.3362	1.3969	1.2174	1.0713
t12	1.0942	1.2605	1.4820	1.3444	1.1488	1.3739	1.2103	1.6458	0.8436	1.1332	1.4041	1.3515	1.1471	1.0064	0.8373
t13	1.0584	1.0415	1.2278	1.1116	1.2333	1.1188	1.1478	1.3942	1.0726	1.3118	1.2667	1.0068	1.0665	0.8558	0.7590
t14	1.3899	1.2316	1.3194	1.3298	1.2384	1.3463	1.6137	1.7244	1.2131	1.5718	1.5810	1.3527	1.2803	1.0654	0.9631
t15	1.0516	1.0759	1.3125	1.1536	1.1014	1.1692	1.1749	1.4437	0.9795	1.2570	1.2670	1.1216	1.0320	0.8292	0.6951

A matrix with distances for calculating the Coarticulation Resistance Coefficient (CRC) of V2 for subject S2 (cf. Section 7.3.2). The labels in the first column represent the 15 $a2_t$ curves, and the labels in the first row represent the 15 $a2_k$ curves. In total, 225 average nearest neighbour distance values are obtained. Each number represents the absolute difference between the two curves concerned, using the Nearest Neighbour technique (see Sections 3.4.6 and 3.4.7.1).

APPENDIX VII. Poster presented at the 5th International Conference on Speech Motor Control, Nijmegen, June 2006



Studying Coarticulation Resistance with Ultrasound

Natalia Zharkova & Nigel Hewlett
Queen Margaret University College, Edinburgh, UK
SSRC Speech Science Research Centre

BACKGROUND

The notion of Coarticulation Resistance (see the Degree of Articulatory Constraint model, Recasens et al. 1997) is implicated in e.g. each of the following claims:

- Vowels are less influenced by consonants than consonants by vowels (Keating et al. 1994)
- Prosodic factors influence segmental coarticulation (e.g. Recasens 2002)
- CV coarticulation is stronger in velars than in alveolars (Modarresi et al. 2004)

Aim: to use ultrasound to study coarticulation resistance in V_1#CV_1 sequences

HYPOTHESES

1. V-on-C effect will occur in V_1#tV_1 sequences *(resistance of /t/ to V is not absolute)*
2. C-on-V effect will occur in a#Ca sequences *(resistance of /a/ to C is not absolute)*
3. V-on-C effect will be stronger than C-on-V effect *(V is more resistant than C)*
4. V2-on-C effect will be stronger than V1-on-C effect *(C is less resistant to V2 than to V1)*
5. V2-on-/k/ effect will be stronger than V2-on-/t/ effect *(/k/ is less resistant to V2 than /t/)*

METHOD

QMUC ultrasound system:

4 COMPARISON OF TWO CURVE SETS

→ *Across Group protocol:*

- Comparing each red curve to each black curve
- 225 distance values generated
- Mean distance value obtained (= size of effect)

→ *Within Group protocol:*

- Comparing each curve in each set to all the other curves from the same set
- 2 sets of 105 distance values generated
- 2 mean distance values obtained, 1 for red and 1 for black

→ *Defining whether the effect is significant:*

- Across Group protocol + 2 Within Group protocols, 1 for red and 1 for black
- Statistical comparison of the Across Group set of distances to each of the Within Group sets of distances
- Result: if both differences are significant then the 2 sets of curves are different

RESULTS

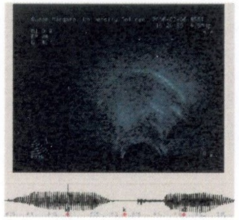
1. Significant V-on-C effect **2. Significant C-on-V1 and C-on-V2 effects**

In all these three graphs, the two sets of curves are significantly different from each other!

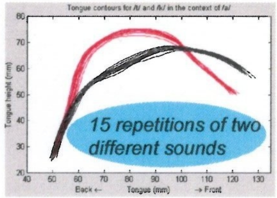
- Three speakers of Southern British English
- Data: /a#ka/, /a#ta/, /i#ti/ in meaningful sentences (e.g. "After that Ma cast an angry look at Leigh")
- Fifteen repetitions

Analysis

1
- Three time points: V1, C, V2
- Drawing curves at time points


mid V1 ⟵ mid /k/ ⟶ same distance /k/ - V2

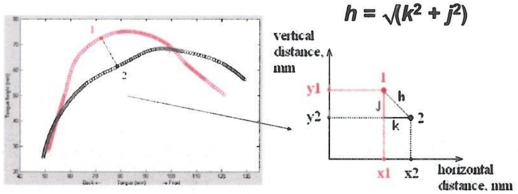
2 xy data for each curve imported into Matlab for plotting and analysis



3 COMPARISON OF TWO CURVES

Finding the shortest distance from each point on one curve to the other curve - nearest neighbour distance:

$$h = \sqrt{(k^2 + j^2)}$$



Then averaging all the nearest neighbour distances, resulting in one average nearest neighbour distance value for the two curves.

3. V-on-C effect significantly stronger than C-on-V effect

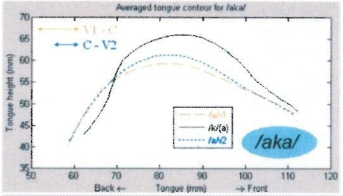
Subject	V-on-C	C-on-V1	C-on-V2
S1	4.51	1.42	1.49
S2	4.48	1.19	1.34
S3	4.26	1.12	1.84

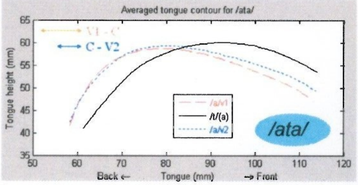
Average distances, mm

Significantly greater average distances for V-on-C effect than for C-on-V1 and C-on-V2!

4. V2-on-C effect significantly stronger than V1-on-C effect

C-V2 average distances significantly smaller than V1-C average distances!





REFERENCES

- Keating, P., Lindblom, B., Lubker, J., & Kreiman, J. (1994). Variability in jaw height for segments in English and Swedish VCVs. *Journal of Phonetics*, **22**, 407-422.
- Modarresi, G., Sussman, H., Lindblom, B. & Burlingame, E. (2004). An acoustic analysis of the bidirectionality of coarticulation in VCV utterances. *Journal of Phonetics*, **32**, 291-312.
- Recasens, D., Pallares, M. D., & Fontdevila, J. (1997). A model of lingual coarticulation based on articulatory constraints. *JASA*, **102**, 544-561.
- Recasens, D. (2002). An EMA study of VCV coarticulatory direction. *JASA*, **111 (6)**, 2828-2841.

5. V2-on-/k/ effect NOT stronger than V2-on-/t/ effect

If we take V1-C distance as 100%, then we can compare C-V2 distance in /aka/ vs /ata/ *(see graphs on the left)*

Subject	C-V2 as percentage of V1-C	
	aka	ata
S1	95.33%	84.29%
S2	84.20%	82.04%
S3	83.78%	82.19%

- *In one subject (S1), V2-on-/k/ effect is significantly weaker than V2-on-/t/ effect!*

(because C-V2 distance, as related to V1-C distance, is greater in /aka/ than in /ata/)

- *In two subjects, no significant differences!*

CONCLUSIONS

- Both C and V are coarticulated, i.e. not maximally resistant to neighbouring segments in VCV
- V more resistant than C to neighbouring segments in VCV
- C less resistant to V2, the vowel belonging to the same syllable of the VCV, than to V1
- /k/ is not less resistant than /t/ to V2, the vowel belonging to the same syllable of the VCV
- DAC values can only be specified in relation to contextual segmental and prosodic interaction (e.g. V or C, syllable structure, possibly more factors not mentioned in our study)
- Individual variation may affect DAC values
- DAC values could be formulated more accurately with this method than has hitherto been offered

ACKNOWLEDGEMENTS: we are thankful for technical assistance to Yolanda Vazquez Alvarez, Steve Cowen and Dr Alan Wrench, and for advice on statistical analysis to Dr Robert Rush, Centre for Integrated Healthcare Research, QMUC