PRODUCTION, PERCEPTION, AND EMERGENT PHONOTACTIC PATTERNS: A CASE OF CONTRASTIVE PALATALIZATION

by

Alexei Kochetov

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Linguistics
University of Toronto

0-612-63688-7

Production, perception, and emergent phonotactic patterns: A case of contrastive palatalization

Doctor of Philosophy
Alexei Kochetov
Department of Linguistics
University of Toronto

The goal of this thesis is to answer two related questions: (i) to what extent can limitations of human speech production and perception explain cross-linguistic positional markedness asymmetries? and (ii) is it logically necessary to attribute these scales to Universal Grammar, as is commonly assumed (e.g., Prince & Smolensky 1993)?

In order to answer these questions, a case study involving the distribution of the plain-palatalized contrast in labial and coronal stops was carried out. A typological survey of languages with contrastive palatalization shows that the distinction between plain and palatalized segments is most often maintained in the syllable onset position and most commonly neutralized in the preconsonantal coda environment. Palatalized labials are more susceptible to neutralization than the palatalized coronals. The most common outcome of neutralization is a plain segment.

In a number of articulatory, acoustic, and perceptual experiments I investigate Russian plain and palatalized stops in cross word-boundary sequences. The articulatory study reveals that the environments that generally induce neutralization exhibit the most variability in the magnitude and timing of the tongue body gesture, the articulatory correlate of the plain-palatalized contrast. They also show that that the effect of environment is different for palatalized labial and coronals. In addition, the variable overlap of primary gestures has important acoustic consequences: it results in the lack of acoustic release burst of the first consonant and less distinct vocalic transitions.

The perceptual findings from both native and non-native subjects under several conditions demonstrate that listeners reliably distinguish the contrast in the contexts when the respective gestures are stable (as in syllable onset), and fail to hear it in the environments that induce gestural variability (syllable coda, especially before consonants).

The results of perception, taken as derived scales, serve as input to a hypothetical learner constructing language-particular grammars. Crucially, the learner is not equipped with positional markedness scales. Limitations on what can or cannot be recovered severely restrict the learning path, ultimately resulting in a limited set of possible grammars that correspond to the attested cross-linguistic patterns of contrastive palatalization.

Acknowledgements

This thesis could not have been possible without the support of many individuals. First and foremost, I wish to express my deep gratitude to Keren Rice, the best thesis advisor and teacher I could have wished for. Her constant encouragement and advice since my first year in the department fostered my interest in phonology, kept me focused and enthusiastic about the topic, and helped me in getting things done. Her always amazingly prompt and insightful feedback was instrumental in the completion of this thesis.

I also wish to thank the other members of my thesis committee — Elan Dresher, Hank Rogers, Ron Smyth, and Joseph Schallert — for their time, effort, and exceptional advice on the various aspects of this research project, not to mention their help with the style and grammar of the thesis.

I would like to express appreciation to Louis Goldstein for his crucial assistance with the articulatory experiments at Haskins Laboratories, New Haven, CT. I have benefited greatly from our many discussions and from his inspirational course on Articulatory Phonology at Yale University. I wish to thank Elizabeth Hume, who acted as an external appraiser, for many insightful comments and suggestions. Bill Idsardi provided invaluable advice at the earlier stages of the thesis. I have also benefited from discussions of parts of the work at meeting of the University of Toronto Phonology Project, as well as from various comments and questions from Peter Avery, Christina Bethin, Ed Burstynsky, Adamantios Gafos, Daniel Hall, David Odden, John Ohala, Janet Pierrehumbert, Glyne Piggott, Charles Reiss, Milan Rezac, Daniel Silverman, Donca Steriade, Caro Struijke, and Lisa Zsiga, among others.

My doctoral studies, experiments, fieldwork, and conference presentations would not have been possible without fellowship and grant funding from the Social Sciences and Humanities Research Council of Canada, the University of Toronto School of Graduate Studies, and the Department Graduate Research Grants to Keren Rice, Elan Dresher, and Hank Rogers. I am also grateful to the many individuals who provided invaluable assistance of various kinds with experiments and presentation of the results: Larissa Chen, Nila Friedberg, Linda Heijmans, Catherine Kitto, Maho Kobayashi, Marianne Pouplier, Michelle Sancier, Mark Tiede, Jeff Weihing, and Patrick Wong. I appreciate the assistance I received from Irina Rusinova and Tatyana Permyakova of Perm' University during my fieldtrip to the northern Urals, Russia. Many thanks to Jack Chambers for helping me materialize some of the results of the fieldtrip.

I would like to show appreciation to all fellow graduate students who shared with me the ups and downs of the graduate life. In particular, I am thankful to Luba Butska, Juli Cebrian, Naomi Cull, Nila Friedberg, Carrie Gillon, Catherine Kitto, Nicole Rosen, and Carolyn Smallwood for many great parties, outings, and just fun. *Spasibo*! Thanks also to my extra-linguistic friends, Jeff and Eileen Anderson, Vladimir Smetanin and Debbie Ehrlich, Tom and Mary LeClerc, Andriy Luchkovsky and Atsuko Takebayashi, Li-Lin Tseng, and to my friends from ISM and JCSA.

Moreover, I am indebted to my parents Rudolf and Lydia and to my sister Elena, who throughout my life have always been there to offer encouragement and support.

Finally, I am grateful to my wife, Maho, for her love, laughter, and patience. This thesis is dedicated to her.

真帆へ

Table of contents

Detailed table of contents

Chapter 1. Foundations

1.0 Introduction

In an examination of phonotactic patterns, it is striking how recurrent such patterns are cross-linguistically. For instance, Steriade 1997 examines the distribution of laryngeal contrasts in languages, finding that certain positional environments systematically support the contrast between voiceless and voiced obstruents, while others tend to neutralize it. An extensive survey of morpheme-internal consonant clusters in Australian languages (Hamilton 1996) reveals that the distribution of place and manner distinctions in these languages is highly restrictive and shows a limited number of patterns. In this dissertation, I enrich the database of phonotactic studies with an examination of the distribution of the plain-palatalized contrast in stops.

The dominant paradigm in phonological theory today, Optimality Theory (Prince & Smolensky 1993), holds that all phonological similarities between human languages are attributable to a substantive component of the grammar, a set of universal markedness constraints. These violable constraints directly state cross-linguistically preferred and disfavoured phonotactic patterns. For instance, the fact that the contrast between voiceless and voiced obstruents is often neutralized in syllable coda position but not in syllable onset environment is assumed to be directly stated in Universal Grammar as harmonically ordered constraints, or a markedness scale: *Voiced-Coda » *Voiced-Onset. This means that a voiced obstruent is more marked in coda position than in onset position. Thus, we are likely to find that if languages maintain the voice contrast in codas they also preserve it in onsets. The opposite does not hold. These markedness constraints, although often assumed to be grounded in phonetic properties (e.g. Archangeli & Pulleyblank 1994), or even directly encoding phonetic properties (e.g. Steriade 1997), are ultimately seen as a part of Universal Grammar (UG), the implicit knowledge that guides a learner in language acquisition.

In this thesis, I present an alternative view of phonotactic patterns. I argue that the set of possible phonotactic patterns is not derived from UG alone, but results largely from an interaction of cognitive phonological properties together with phonetic factors (articulatory, acoustic, perceptual). Some of these factors, as I show, are external to the phonological grammar proper.

Below I outline the assumptions made in the thesis about articulatory gestures and the relation between production and perception, and present a model of neutralization of phonological contrasts.

1.1 Articulation gestures, production, and perception

In this thesis I assume the framework of *Articulatory Phonology* (Browman & Goldstein 1986 et seq.). The key concept of this theory is an *articulatory gesture* (Liberman & Mattingly 1985), which is assumed to have a dual nature. First, it is a unit of articulatory motor action, or the coordinated activity of articulators, such the lower and upper lips achieving a closure. Second, it is a linguistic unit of information that distinguishes phonological contrasts, such as [labial] and [closed]. While the gesture as unit of action is gradient, physical, and concrete, the gesture as a phonological unit is assumed to be on the categorical, cognitive, and abstract side of the spectrum. To avoid confusion between the two units I refer to the motor unit simply as a *gesture* and to the phonological units as a *gesture/feature*. It is important, however, that the two units are intricately linked representing two different sides of the same phenomenon, or two different levels of representation that vary in the degree of abstractness (the microscopic and macroscopic levels).[1] The innate gestures/features are considered to be phonological primitives whose stable combinations result in higher level phonological structures, such as segments and syllables (cf. the innate nature of features in Chomsky & Halle 1968).

[1] This view of the relation between phonetics and phonology as a continuum between physical and categorical properties has become prevalent in the field of Laboratory Phonology (cf. Pierrehumbert, Beckman, & Ladd 2001).

Proponents of Articulatory Phonology have argued that reference to the dual nature of the articulatory gesture provides an explanatory account of a range of phonetic and phonological phenomena. An increasing number of studies has been devoted to gestural dynamics (gestural coordination, overlap, reduction, and timing) and its role in phonological assimilation and deletion (Browman & Goldstein 1990, Byrd 1992), allomorphy (e.g. Zsiga 1995), and syllable organization (e.g. Browman & Goldstein 1995, Gick 1999).

The recent interest in Articulatory Phonology involves investigation of the effect of gestural organization on perception of phonological units, or *recoverability of gestures/features*. The relation between gestural coordination and recoverability has been investigated in works by Byrd 1992, 1996, Surprenant & Goldstein 1998, and Chitoran, Goldstein, & Byrd 2000. This research strategy views articulatory gestures/features as self-organizing structures. Stable coordination patterns of gestures/features, commonly found cross-linguistically, evolve due to simple conditions, the most critical of which is the fact that there are limits on the ability of a speaker to recover gestures from a signal (Browman & Goldstein 1998). Thus, recoverability plays the role of a 'filter', selecting more stable gestural structures (or gestural coordination patterns) over less recoverable ones.

In this thesis I extend the hypothesis of gestural/featural recoverability to the phenomena of phonotactic distribution and neutralization of a phonological contrast. In order to provide background for the approach, I outline the basic assumptions of recoverability.

The general model of speaker-listener interaction, or gestural model of perception (based on the *Revised Motor Theory* by Liberman & Mattingly 1985 and *Direct Realist Theory* by Fowler 1986), is presented in Figure 1.1. A motor articulatory movement initiated by a speaker results in an acoustic speech signal. A listener uses this signal as information about the 'underlying' gestures/features of the message. In other words, speech perception is understood as recovering, or interpreting, articulatory gestures/features (as viewed in Articulatory Phonology).

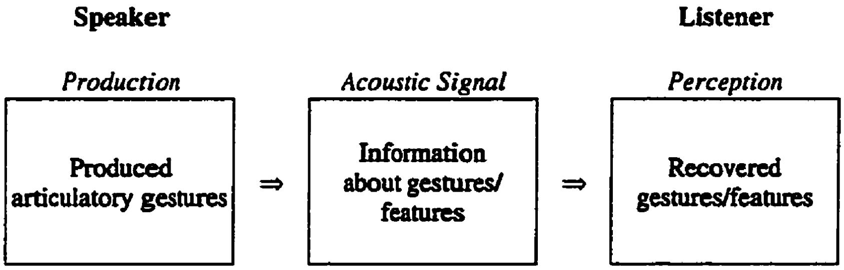


Figure 1.1 Relation between speech production (Speaker) and perception (Listener)

It is important to point out several crucial characteristics of this process. First, according to the model the listener recovers relatively abstract speaker-independent units (i.e. gestures/features) rather than the detailed motor program and vocal tract configuration of the actual speaker (i.e. articulatory gestures per se) (Fowler 1996). Second, the acoustic signal is a lawful consequence of an articulatory movement; however, the relation between the two is non-linear, or one-to-many (Liberman et al. 1967). In some cases minor articulatory changes may have major acoustic consequences (e.g. in stop bursts) and certain different vocal tract configurations may result in similar acoustic outputs (e.g. in production of vowels). Thus the recovery of gestures/features from an acoustic signal is not a straightforward task. Third, the model does not underestimate the importance of acoustic signal in perception. The listener has to have the detailed knowledge of acoustic parameters in order to extract necessary information. However, unlike some other theories of speech perception (e.g. Bladon 1986, Bregman 1990, Johnson et al. 1993, Kingston & Diehl 1994), the model does not assume that the symbolic units we perceive are acoustic in nature. This is, in fact, in agreement with the generally assumed view of features in phonological theory as having an articulatory rather than an acoustic basis (e.g. Chomsky & Halle 1968, Clements 1985, Avery & Rice 1989, Clements & Hume 1995).[2]

Note that it is beyond the scope of this thesis to argue for the gestural model of speech perception, although some findings may be later interpreted in support of it or against it. The goals of the thesis are different. By incorporating this approach into a *model of neutralization* of phonological contrasts I will determine the sources (whether articulatory, acoustic, perceptual) of neutralization of the plain-palatalized contrast. This will provide us with answers to more general questions of the nature of knowledge of phonotactics and what must or may not be present in Universal Grammar. In this respect the work continues the general line of research investigating *phonetic* factors in the search for explanations of *phonological* universals (e.g. Ohala 1981, 1983, Kawasaki 1982, Lindblom, MacNeilage, & Studdert-Kennedy 1983, Maddieson 1984, Silverman 1997, Steriade 1997, Hume & Johnson 1999, Hume 2001, Gafos 2000, among others).

1.2 Neutralization

In the thesis neutralization is understood as a suspension of a phonological contrast between two segments in a given environment (Trubetzkoy 1958/69: 78-83). These segments usually differ in the presence or absence of a gesture/feature, or its different values. Only one of the two segments is found in the neutralizing context. Note that this definition is not limited to alternating forms, so called 'active' neutralization, e.g. final obstruent voice neutralization in German or Russian (/rod/ [rɔt] 'gender' and /rot/ [rɔt] 'mouth'), but also encompasses multiple 'static' cases, e.g. obstruent voice neutralization after /s/ in English (/spɪn/ but */sbɪn/).

In the proposed model of neutralization gestural/featural recoverability plays a crucial role. This is shown in Figure 1.2, which expands Figure 1.1. In 1.2 further information is provided about articulatory gestures/features (presence or absence), the acoustic signal (more or less information), and perception (recovered presence or absence of gestures/features). Production, acoustics, and perception are complemented by the listener's lexicon and grammar. Note that our listener is also a learner that is acquiring language phonotactics.

[2] But see Jakobson, Fant & Halle 1963, Flemming 1995, and Hamilton 1996 for arguments for acoustically based features, or a combination of both.

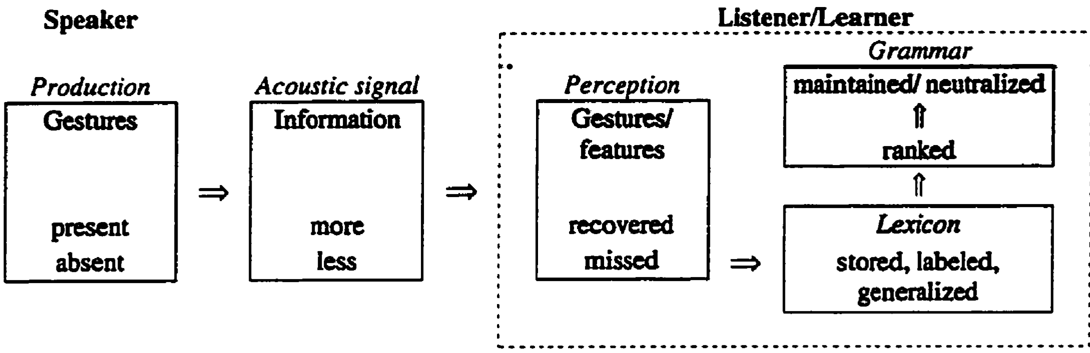


Figure 1.2 A model of neutralization

In order to perceive speech, the listener uses the acoustic signal to extract information about gestures/features produced by the speaker. The success of the listener in recovering these gestures/features, and thus correctly interpreting the message, depends on a number of factors, among which *articulatory* and *acoustic* factors are the focus of this work.

Imagine a hypothetical phonological contrast between two segments, /X/ and /Y/. While the production of segment /X/ involves one articulatory gesture, [G1], the articulation of /Y/ involves two gestures, [G1] and [G2]. Both segments share [G1] and differ in the presence or absence of [G2]. A speaker pronounces two utterances with segments /X/ and /Y/ in the same environment. The articulation of the gestures ([G1] for /X/ and [G1], [G2] for /Y/) results in an acoustic signal whose spectral and temporal properties refer to the intended gestures/features. A listener/learner recovers these units and thus perceives the two segments (and the utterances) as distinct. Thus, the gestural input, articulation, equals the output, perception, presenting a case of perfect recoverability. As a result, the contrast distinguished by the speaker is also distinguished by the listener (/X/ vs. /Y/ = /X/ vs. /Y/).

It is important, however, that articulatory gestures have their own intrinsic properties (duration, constriction location, and constriction degree) and that they overlap with each other in time. Thus, if [G1] is produced simultaneously with [G2] (the common phenomenon of co-articulation), the acoustic information about one of them (e.g. [G2]) may be 'hidden' by the other (e.g. [G1]), depending on their properties. The listener would recover [G1] and would miss [G2].

The consequence of this is that the second utterance (with [G1], [G2]) is not perceived as different from [G1]. As a result of this, the contrast distinguished by the speaker is not distinguished by the listener (/X/ vs. /Y/ = /X/).

This is only one possible scenario. The recoverability of [G2] can also be affected by other co-produced gestures and their acoustic consequences, gestural reduction in certain positions and under certain speaking conditions, etc. Other factors such as non-linear relations between articulation and acoustics and general auditory limitations also contribute to the overall performance of the listener. As a result, some gestural/featural structures (or segments) are more likely to be confused by the listener, and thus neutralized with others. The resulting recoverability scale (e.g. /X/ > /Y/, or /X/ is easier to recover than /Y/), and its modifications depending on various contexts, are a consequence of multiple factors (cf. Hume & Johnson 2001). Its general properties are partly extra-linguistic, shared with other kinds of human perception, e.g. visual or tactile (as also pointed out to me by Elan Dresher).

To summarize, the more distinct the acoustic consequences of the corresponding articulatory movements are, the more likely the listener/learner is to recover the gestures/features, and thus to perceive the utterances as different. On the other hand, subtle acoustic differences (possibly as a consequence minor gestural differences) result in a situation where the listener/learner fails to recover some of the input gestures/features, thus perceiving the utterances as nondistinct.

The last two components of the model, the grammar and the lexicon, are equally important. Lexical items, whether their gestures/features are recovered, missed, or confused, are further processed by the listener/learner in a cognitive mode (i.e. pertaining to the listener's cognitive system) with reference to already acquired lexical and phonological categories of the language. Thus, a perceived difference or similarity in a given context may be generalized in one or the other direction, depending on the already established lexicon, and consequently, the contrast is maintained or neutralized. The likelihood of neutralization is higher in the case of minor articulatory and acoustic differences between sounds as these are more likely to induce listener errors.

In this model, at any point in time throughout the process, our listener/learner does not have to 'know' that certain gestures/features or segments in certain environments are marked, and thus should be maintained or neutralized. The learner, on recovering the input, simply fails to hear the acoustic consequence of a gesture (and ultimately, gesture/feature or segment) and makes further decisions based on this. Thus the motivation for neutralization has very little to do with the listener/learner's knowledge of a markedness scale and the corresponding ranking of constraints in the grammar. It arises due to the intrinsic nature of articulatory gestures and general principles underlying human perception being modified by the learner's grammar.

Unlike our learner, the assumed OT-type learner is equipped with a markedness scale (e.g. *Voiced-Coda » *Voiced-Onset) that is intended to guide her/him in a decision as to whether to maintain or neutralize a contrast. Like our listener, s/he is likely to fail to recover the relevant gesture under the same circumstances, and to conclude that the contrast is neutralized (cf. Hale & Reiss 2000). Thus, neutralization happens independently of the OT learner's having reference to the positional markedness scale. Note that similar arguments can be presented against the view which proposes that the listener/learner is guided by knowledge of the phonetic cues to a contrast and their relative salience in different environments, a view argued for by, for instance, Steriade 1997, Jun 1995, and Flemming 1995.

In sum, the model of neutralization of a positional contrast is crucially based on the role of gestural recoverability, being an interaction of articulatory, acoustic, and perceptual factors with cognitive phonological (and lexical) effects. An important consequence of assuming this model is a view of UG with an impoverished substantive component and thus simpler and computationally more plausible language-particular grammars.

1.3 Hypothesis and experiments

The model outlined in the previous sections makes a number of straightforward predictions about neutralization environments that can be experimentally verified. These are given in general form below.

If a phonological contrast in languages shows recurrent patterns of neutralization of the contrast in one environment but not in another, the motivation for the pattern is in asymmetries in the recoverability of the contrast (i.e. corresponding gestures/features) in one environment compared to the other, due to certain articulatory and acoustic factors. Also, if some contrast shows higher susceptibility to neutralization than another in the same context, this may be due to the relatively less reliable recoverability of at least one of these segments as a consequence of smaller articulatory and acoustic differences. The differences between the same gestures in different contexts and different gestures in the same context may be a product of a number of independent factors. Thus in order to determine the exact factors and their interaction one has to examine articulatory and acoustic properties of the contrast in a variety of contexts and to test how they are perceived by listeners under several conditions. Finally, one should have a testable hypothesis about how these factors directly interact with the phonological grammar.

All these issues are addressed and the corresponding predictions about neutralization are experimentally tested in the thesis. The case study presented here involves the distribution of the plain-palatalized contrast in labial and coronal stops, focusing on Russian. I examine the cross-linguistic patterns in languages that exhibit the contrast and identify the most and least likely neutralization sites and the segments that are more or less susceptible to neutralization. I further investigate the articulatory and acoustic properties of plain and palatalized segments in a variety of contexts and test the recoverability of the corresponding gestures/features. Further, the recoverability results, taken as derived perceptual scales, serve as input to a hypothetical learner constructing language-particular grammars. This produces a range of possible lexicons and grammars that are compared to the attested ones. Thus the model is tested at each stage of the process, production, acoustic signal, perception, and grammar (see Figure 1.2). An examination of one component makes additional predictions for other components.

The thesis is organized on a series of experiments, which correspond roughly to the components of the model. Chapter 2 examines the phonotactic patterns of contrastive palatalization attested in a number of languages. The pattern exhibited by Standard Russian is investigated in detail, considering frequency and morphological restrictions. I then turn to acoustics (Chapter 3) to investigate the acoustic properties of the transitions and stop release bursts of Russian palatalized coronal stops in coda position, as they are important in later chapters. This is followed by a perceptual experiment testing the importance of transitions and bursts (Chapter 4). Chapter 5 presents an articulatory study of gestures involved in the production of plain and palatalized coronal and labial stops, identifying the factors likely to affect the recoverability of the contrast. The predictions made in this chapter are examined in a series of perceptual experiments in Chapter 6. In Chapter 7 I compare the articulation and perception results and derive perceptual scales as a combined result of articulatory, acoustic, and perceptual factors. Chapter 8 summarizes all the factors and demonstrates how their interaction with phonological factors results in the attested phonotactic patterns.

Chapter 2. Phonotactic patterns of palatalization

2.0 Introduction

As discussed in the introduction, one could imagine that languages show a range of possibilities with respect to phonotactic patterns. They could differ widely in terms of within-language generalizations and could exhibit no cross-language generalizations. The evidence from both domains suggests that the universal nature of phonotactic patterns arises due to the same set of physical and cognitive factors (see Jakobson, Fant, & Halle 1969, Chomsky & Halle 1968, Ohala 1981, 1983, Kawasaki 1982, among others, for different views on the subject).

In this chapter I examine this issue with respect to the distribution of the plain-palatalized contrast in consonants. I show that cross-linguistically, strong patterns of asymmetries are found in the distribution of this contrast with respect to place of articulation and environments. Palatalized coronals are generally less restricted in their distribution than palatalized labials. Overall, the contrast between plain and palatalized consonants is often maintained in syllable onset position, and is more commonly disfavoured in coda, where its occurrence depends on more specific contexts, such as word-final or preconsonantal. I show that a detailed analysis of Russian phonotactics provides a deeper insight into existing asymmetries in the distribution of plain and palatalized consonants by making additional reference to frequency and morphological factors. The findings of this chapter set up the general direction for the focus of the further phonetic experiments (Chapters 3-7). Note that the discussion of phonotactics is descriptive and theory-neutral. An explanatory analysis of the data, based on the results of the experiments, is provided in Chapter 8.

In section 2.1 I present the results of a survey of languages that have contrastive palatalization, examining both positional and featural restrictions that affect the contrast. The results of the cross-linguistic survey are followed by a more detailed account of Russian phonotactics in the selected environments in section 2.2.

2.1 Cross-linguistic patterns of contrastive palatalization

2.1.0 Introduction

This section examines the restriction on the plain-palatalized contrast found in a number of related and unrelated languages. The focus is plain and palatalized labial and coronal stops. The literature survey of languages with contrastive palatalization is followed by general background (phonological inventories and phonetic realization) on these consonants. The main part of this section is the discussion of common phonotactic patterns with respect to palatalization.

2.1.1 The survey

The goal of the survey is to determine the common phonological and phonetic properties shared by languages with a plain-palatalized contrasts, and, more particularly, to examine the restrictions that apply to the distribution of the contrast. The survey is largely restricted to Slavic, Celtic, and Uralic languages. Table 2.1 lists the languages and dialects examined by their language group or family and provides sources. Only languages that are described as having a plain-palatalized distinction (in some form, as discussed later) were selected for analysis. Thus I included East and West Slavic languages and Bulgarian (South Slavic), languages with a contrast in terms of palatalization in both labials and coronals (Carlton 1990).[3] Data from dialects of Bulgarian, Russian, and Polish are included in cases where dialects differ from the standard in relevant features (realization of segments or distribution). In the same way, the languages of other families that are not reported as having (or having ever had) a plain-palatalized contrast at

[3] Not all of these languages retain palatalization in labials, limiting the contrast to coronals only (e.g. Ukrainian). Czech and Slovak present more opaque cases, since they have developed the palatalized coronals into 'true palatal' consonants, while still exhibiting alternations with plain segments (Kuznetsova 1969, Short 1993ab). I did not include the West Slavic languages Kashubian and Upper and Lower Sorbian, where the original coronal palatalized stops/affricates either merged with dental/alveolar palato-alveolar affricates (Kashubian: Stone 1993; Upper Sorbian: Carlton 1990: 254-264) or underwent deaffrication (Lower Sorbian: Carlton 1990: 265-273).

some place of articulation are not included here (e.g. Latvian (Baltic) (Chekman 1970); the Celtic languages Breton and Welsh (Ball & Fife 1993); the Uralic languages Finnish, Izhora, Vod', Hungarian, Mari, Udmurt, Komi, Khanty, Mansi (Abondolo 1998, Iartseva 1993)).[4] For some of the languages with reported palatalized consonants (e.g. the Uralic languages Estonian, Moksha Mordva, Veps, Liv, and Saami (Abondolo1998, Iartseva 1993), information about their phonotactics is limited or unavailable.

Table 2.1 Languages used in the survey

Group/Family	*Language*	*Sources*
Slavic (East)		
	Belorussian	Mayo 1993, Carlton 1990: 293-301
	Russian	Avanesov 1972, Bondarko 1981, Jones & Ward 1969, Timberlake 1993, Borkovskii & Kuznetsov 1965, Kiparsky 1979, Carlton 1990: 285-292
	Russian dialects	Avanesov 1972, Avanesov & Orlova 1965, Avanesov & Bromlei 1986, Azarkh 1967, 1973, Kuznetsova 1969, 1977, Orlova 1970, Kochetov 1999
	Ukrainian	Shevelov 1993, Carlton 1990: 277-285
Slavic (West)		
	Czech	Short 1993a, Bernshtein 1977, Carlton 1990: 231-238
	Polish	Rothstein 1993, Bethin 1992, Bernshtein 1977, Carlton 1990: 248-253
	Polish dialects	Basara et al. 1959, Dejna 1993, Stieber 1968, Zdunska 1965
	Slovak	Short 1993b, Bernshtein 1977, Carlton 1990: 238-245
Slavic (South)		
	Bulgarian	Scatton 1993, Bernshtein 1977, Carlton 1990: 303-306
	Bulgarian dialects	Chekman 1970, Khristov 1956, Joseph Schallert p.c.
Baltic		
	Lithuanian	Ambrazas 1997, Mathiassen 1996
Celtic		
	Irish	Ball & Fife 1993, De Burca 1958, MacAulay 1992, Russel 1995
	Manx	Ball & Fife 1993, MacAulay 1992
	Scots Gaelic	Ball & Fife 1993, MacAulay 1992, Ternes 1973
Uralic		
	Erzya Mordva	Abondolo 1998, Redei 1984
	Karelian dialects	Leskinen 1968
	Nenets	Abondolo 1998, Decsy 1966
Other		
	Japanese	Vance 1987, Akamatsu 1997, 2000, Wenck 1966

[4] Some of these languages (Latvian, Breton, Hungarian, Udmurt, Komi, among others) have 'true palatals', rather than palatalized dentals/alveolars (Chekman 1970, Abondolo 1998, Iartseva 1993, Lytkin 1962).

2.1.2 Plain and palatalized consonants: General

In this section I examine the general properties of plain and palatalized consonants (phonological contrast and phonetic realization) and provide a brief background on the historical sources of palatalized consonants.

2.1.2.1 Phonological contrast

In languages with a plain-palatalized contrast, consonants of almost any place or manner of articulation can be palatalized.[5] I use the apostrophe sign on a consonant, C', when discussing the phonemic status of palatalized consonants (as opposed to the IPA symbol [C^j] that will be used for the phonetic realization of palatalization).

The consonant inventory of Bulgarian (Scatton 1993: 191), shown in (2.1) is typical of a language with contrastive palatalization: labials, coronals, and velars of various manners of articulation exhibit a contrast. It is common that consonants that are not paired with respect to palatalization (like the velar fricative /x/ and post-alveolars) pattern phonologically together with either plain or palatalized segments.

(2.1) Bulgarian phonemic consonant inventory

[illegible]		[illegible]			[illegible]	
plain	pal	plain	pal	pal	plain	pal
p	p'	t	t'		k	k'
b	b'	d	d'		g	g'
		ts	ts'	tʃ		
		(dz)	(dz')	dʒ		
f	f'	s	s'	ʃ	x	--
v	v'	z	z'	ʒ		
m	m'	n	n'			
		l	l'			
		r	r'			
				j		

I focus on the distribution of plain and palatalized labial and coronal voiceless stops (2.2).

[5] Retroflex and uvular consonants are likely to be exceptions (based on Maddieson 1984).

(2.2) a. /p/ vs. /p'/

b. /t/ vs. /t'/

There are several reasons, both conceptual and methodological, for this choice of consonants. First, palatalized labials and coronals are more commonly found in the surveyed languages than palatalized velars.[6] Second, labials and coronals show a number of asymmetries with respect to their distribution, while palatalized velars (at least in the Slavic languages that have them) are often limited to the syllable onset (prevocalic) position and are marginally contrastive before vowels (mainly back vowels). Third, the articulatory properties of palatalized labials and coronals can be easily observed in an articulatory magnetometer study, since their primary (lips and tongue tip) and secondary (tongue body) articulators are relatively far apart from each other. The articulation of both plain and palatalized velars involves movement of the tongue dorsum, and this strongly affects the trajectory of the tongue body. Fourth, the choice of stops rather than fricatives or nasals is motivated by the fact that stops are usually better documented in the phonological and phonetic literature. This is also true of voiceless stops compared to voiced ones.

2.1.2.2 Phonetic realization

The four phonemes /p/, /p'/, /t/, and /t'/ can be realized phonetically in the surveyed languages in a variety of ways, as shown in Table 2.2.

The most common realization of the /p/ vs. /p'/ contrast is the simple absence or presence of the secondary palatal articulation: [p] and [p^j]. In some cases, however, the plain consonants can also be velarized (e.g. in varieties of Irish and Russian) or labialized (other varieties of Irish). In fact, at least in some dialects of Irish the contrast is often velarized/labialized vs. plain, since the phonemic palatalized labial rarely has any secondary palatal articulation (De Burca 1958,

[6] This gap in the Slavic languages can be partly explained by a number of Common Slavic changes of velars in the environments of front vowel and yod to palatal or post-alveolar affricates and fricatives (palatalization processes: Carlton 1990: 112-116, 120-126).

MacAulay 1992). Scots Gaelic is another example of an opaque /p/ vs. /p'/ contrast, where the phonemic palatalized segment differs from /p/ only by lip spreading (Ball & Fife 1993: 154) or is phonetically identical to it (MacAulay 1992: 231, Henry Rogers, p.c.). Given this, it is not surprising that the phonological status of /p'/ is disputed (Ternes 1973: 32-52).

Table 2.2 Realizations of plain and palatalized voiceless labial and coronal stops

	/p/	/p'/	/t/	/t'/
Belorussian	p	pʲ	t̪	tsʲ
Bulgarian	p	pʲ	t	tʲ
Bulgarian dialects	p	pʲ / --	t	tʲ / c
Czech	no contrast		t	c
Erzya Mordva	no contrast		t̪	tʲ
Irish	pˠ / pʷ	p	t̪ˠ / t̪ʷ	tɕ / tʃ / t
Japanese	p	pʲ	t	tɕ
Karelian	no contrast		t	tʲ
Lithuanian	p	pʲ	t̪	tʲ / tʃ
Manx	no contrast		t̪	tʃ
Nenets	p	pʲ	t	tʲ
Polish	p	pʲ	t̪	tɕ
Polish dialects	no contrast		t̪	tɕ / tʃ
Russian	pˠ	pʲ	t̪ˠ	tʲ
Russian dialects	p	pʲ	t / t̪	tʲ / tɕ / c / tsʲ
Scots Gaelic	p	p	t̪	tɕ / tʃ
Slovak	no contrast		t	c
Ukrainian	no contrast		t̪	tʲ

Notes: In Irish stops are aspirated;
In Scots Gaelic /p'/ is realized as [p] or with additional spreading of lips; stops are pre- or post-aspirated (Ball & Fife 1993: 154-155);
In addition to /t'/ [tɕ], Polish has phonetic palatalized [tʲ] in loans (Bernshtein 1977);
In Lithuanian /t'/ [tʃ] is realized as [tʲ] before front vowels (Ambrazas 1997);
See the discussion of languages with no contrast in labials below.

There are several points to note with regard to the realization of the plain-palatalized contrast in coronals. First, plain stops are usually realized as dental or alveolar consonants. The dental articulation, however, seems to be more common in these languages.[7] Only Irish is noted to have additional velar/labial articulation on the dental [t]. Second, the palatalized coronal /t'/ shows a wide range of realizations ([tʲ], [tʃ], [tɕ], [c]).[8] All of these involve the raising and

[7] It should be noted that the grammars do not always specify the exact place of articulation.
[8] Many languages have additional coronal affricates (alveolars, palato-alveolars, and retroflexes). The variation in realizations of /t'/ appears to be more limited in these cases (e.g. in Slavic languages compared to Celtic languages).

fronting movement of the tongue body to the hard palate; this movement usually results in a laminal shape of the tongue front. The realizations do, however, differ in the primary constriction (tip/lamina at the upper teeth/alveolar ridge for [t^j] or [ts^j], lamina at the post-alveolar region for [tʃ], lamina/front at the prepalatal region for [tɕ], and front of the tongue at the hard palate for [c]). In other words, all these realizations present points on the articulatory continuum between [t^j] and [c], with the first one being a prototypical palatalized dental/alveolar consonant. They also show a continuum in the degree of burst release duration that makes the distinction between 'true' stops and 'true' affricates difficult to make (Kuznetsova 1977).

We see that not all of the languages that have the contrast in coronals (/t/ vs. /t'/) maintain it in labials, having only plain labial /p/ instead. In many of these languages the absence of /p'/ corresponds to the presence of a /p/+/j/ cluster (e.g. Czech, Ukrainian, north-eastern Polish dialects, and Manx). However, the opposite does not hold: some languages may allow both the segment /p'/ and the sequence /pj/ (e.g. Russian (see section 2.2.2.1.2)). Plain /p/ in languages without a /p/ vs. /p'/ contrast may be realized as [p^j] before front vowels (e.g. Ukrainian (Carlton: 283-284)).

2.1.2.3 Summary and conclusion

In sum, there are differences between labials and coronals with respect to their secondary articulation. Palatalized labials are less common in languages than palatalized coronals, and they are more restricted in their realizations. This is reflected in the implicational statements in (2.3).

(2.3) Inventory

a. p vs. p' > t vs. t'
• If a language has a plain-palatalized contrast in labials, it also has it in coronals

b. p' > t'
• If a language has a palatalized labial, it also has a palatalized coronal

The most common output of /p'/ is [p^j], and the prototypical realization of /t'/ is [t^j]. The articulation of the latter, however, is also commonly found further back in the post-alveolar region and it is often characterized by strong affrication.

I will refer to the languages that have the plain-palatalized contrast in both coronals and labials as having a four-way contrast (/p/, /p'/, /t/, and /t'/) and those that limit it to coronals only as maintaining a three-way distinction (/p/, /t/, and /t'/).

2.1.2.4 The direction of sound change

So far I have focused on the synchronic distribution of the plain-palatalized contrast. Given the implications seen in (2.3) we might expect that these would not only hold of phonotactic patterns, but might also be useful in predicting directions of sound change.

In many of the languages in the survey the palatalized consonants are known to have developed due to several processes: the phonetic palatalization of consonants in the environment of front vowels and yod, followed by backing of the following vowel (onset position); apocope (final coda position); and syncope (preconsonantal coda position) (Slavic: Carlton 1990: 157-164, Borkovskii & Kuznetsov 1965: 112-113, 122-124, Kiparsky 1979: 121, 126-128; Celtic: Ball & Fife 1993, Russel 1995).

Many of the three-way contrast languages (e.g. Czech, Slovak, and Manx, among others) originally had the plain-palatalized distinction in labials. The common process of depalatalization involved the change of the final /p'/ to plain /p/ and the initial /p'/ to the sequence /pj/, which led to the elimination of the /p/ vs. /p'/ contrast altogether (Carlton 1990: 160-162; MacAulay 1992). Some of the four-way languages show variation with respect to the realization of the onset /p'/ (e.g. Polish (Bethin 1992), Bulgarian (Scatton 1993: 190)) or additional positional restrictions on it (Lithuanian (Ambrazas 1997)).

The palatalized coronal shows a stronger resistance to depalatalization, at the same time exhibiting a tendency to affrication and a shift in place of articulation to the post-alveolar region (Kuznetsova 1977, Chekman 1970).

2.1.3 Plain and palatalized consonants: Distribution

In the sections below I examine the distribution patterns of /p/, /p'/, /t/, and /t'/ as single consonants and segments in clusters. Since both initial and medial onset consonants in most of these languages do not vary in what segments they allow, I will refer to this environment as the 'onset environment', or 'onset position' (the special cases of Lithuanian and Karelian will be discussed in section 2.1.3.1.2). The onset consonants in clusters, as discussed in section 2.1.3.2.3, also fall into the same class. However, single (final) coda consonants and the same consonants in clusters (preconsonantal) do not always pattern in the same way. I distinguish between the two environments as 'final (coda)' and 'preconsonantal (coda)' (or a 'coda in cluster') and will discuss them separately. The environments are summarized in the following order: single consonants (onset and final coda environments) and consonants in clusters (preconsonantal coda before a plain consonant and preconsonantal coda before a palatalized consonant).

The patterns in sections 2.1.3.1-2.1.3.2 are presented from the least restricted to the most restricted. Section 2.1.3.3 presents a summary. The questions of interest are: which places of articulation are more susceptible to neutralization of the plain-palatalized distinction, what is the common outcome of this neutralization, and what environments induce this process.

2.1.3.1 Single consonants

In this section I consider the distribution patterns of plain and palatalized labials and coronals as single consonants.

2.1.3.1.1 Onset and final (coda) environments

2.1.3.1.1.1 Data

There are three possible patterns attested in the languages with a 4-way plain-palatalized contrast (labial vs. coronal, plain vs. palatalized) and with consonants in final position. In one pattern, all four segments can be found both in onset and final coda without any restriction, as shown in (2.4). Here and below I use nonsense examples (*pa, p'a*, etc.) to illustrate well-formed and ill-formed strings. Note that phonetic realizations of /p/, /p'/, /t/, and /t'/ may vary from language to language (see Table 2.2). The vowel /a/ stands for any back vowel.

(2.4) Pattern 1: Russian, Irish, Bulgarian (Nova Nadezhda and Bela Cherkva)

	p	t	p'	t'
onset	*pa*	*ta*	*p'a*	*t'a*
final coda	*ap*	*at*	*ap'*	*at'*

This pattern is found in Standard Russian, some dialects of eastern Bulgarian (Nova Nadezhda and Bela Cherkva), and Irish. Thus, /pa/, /ta/, /p'a/, and /t'a/ are possible initial sequences in these languages, and /ap/, /at/, /ap'/, and /at'/ are legal final strings. The contrast in terms of palatalization is maintained both in the onset and in the final coda environments.

The second pattern (2.5) is quite common, being attested in many northern, western, and southern Russian dialects (e.g. Cherdyn': Kochetov 1999), Belorussian, Polish, and some dialects of eastern Bulgarian (Krinichnoe: Chekman 1970). It is also reported for Scots Gaelic (assuming the phonemic status of palatalized labials: Ball & Fife 1993).

(2.5) Pattern 2: Russian (Cherdyn'), Belorussian, Bulgarian (Krinichnoe), Scots Gaelic

	p	t	p'	t'
onset	*pa*	*ta*	*p'a*	*t'a*
final coda	*ap*	*at*	[redacted]	*at'*

Note: * = disallowed

In these languages the contrast between the plain and palatalized coronals is possible in both onset and final coda environments; however, the distinction between labials is neutralized finally, where only plain /p/ is found. Thus, /ap'/ is not a phonotactically possible sequence in these languages. In this pattern the contrast is maintained in the onset position for both labials and coronals. In the final coda, it is also retained for coronals, but not for labials.

The third possible pattern exhibits neutralization of the plain-palatalized contrast finally for both labials and coronals (2.6). In both cases only plain /p/ and /t/ are found in this position. Languages like Standard Bulgarian, Lithuanian, Nenets, and some northern Russian dialects (Vologda and Vyatka regions: Azarkh 1967, 1973) exhibit this distribution pattern. Pattern 3 maintains the palatalization contrast in the onset and neutralizes it in the final coda position, regardless of the place of articulation of the consonants.

(2.6) Pattern 3: Russian (Vologda-Vyatka), Bulgarian (Standard)

	p	t	p'	t'
onset	*pa*	*ta*	*p'a*	*t'a*
final coda	*ap*	*at*		

Some of the languages of Patterns 2 and 3 show active neutralization in the form of alternations between final /p/ and prevocalic /p'/ (e.g. Bulgarian (Scatton 1993: 197)).

In Japanese the plain-palatalized distinction is possible in onsets (*pa, p'a, ta, t'a*) but not in final codas (**ap, *ap', *at, *at'*). However, this is due to the more general restriction against all final coda obstruents, *C_{obs}# (Vance 1987). This pattern is also important for our discussion, since it involves an asymmetry between the two environments, initial and final.

Note that none of the examined languages show neutralization of the contrast in coronals but not in labials (i.e. *at vs. *at'*, but *ap vs. ap'*), or restrictions against final plain consonants at the expense of the palatalized ones (i.e. **ap, *at,* but *at', ap'*).

The languages with a three-way plain-palatalized contrast (/t/ vs. /t'/ only) show similar distribution patterns. As in Patterns 1 and 2, the contrast /t/ vs. /t'/ is unrestricted with respect to position in Czech, Erzya Mordva, Manx, Slovak, Ukrainian, and an Oloneck dialect of Karelian (Oloneck 1: Leskinen 1968). In these languages *ta, t'a, at* and *at'* are well-formed sequences. Other languages, like Nenets and other dialects of Karelian (Arkhangelsk and Oloneck 2: Leskinen 1968) neutralize the distinction finally, having sequences of the type *ta, t'a*, and *at*, but not *at'*. This distribution pattern corresponds to Pattern 3. In general, languages with a three-way contrast can be seen as taking the neutralization of the plain-palatalized distinction in labials a step further, to the onset position. No parallel is logically possible with Pattern 2.

To summarize, languages that have a four-way plain-palatalized contrast (and permit final consonants) differ in whether this contrast is allowed both initially/medially and finally. Out of a number of possible patterns, only three are attested. In all of them both plain and palatalized consonants are found in syllable onsets (initial or medial). The patterns differ in whether palatalized consonants are permitted in final coda position. If there is a restriction against them, it is primarily applied to the palatalized labial /p'/.

2.1.3.1.1.2 Summary and generalizations

The neutralization patterns discussed above are summarized in (2.7), where onset and final coda coronals and labials are given in separate columns. 'Yes' indicates that a particular contrast is maintained in a given position, while 'no' denotes its neutralization. The segment in parentheses shows the outcome of the neutralization. Thus, Pattern 1 allows the contrast in both labials and coronals in both of the environments. Pattern 2 is more restrictive: the final labials are neutralized in favour of the plain /p/. Pattern 3 shows neutralization of both labials and coronals finally, with the plain segments as default outcomes. As I mentioned earlier, languages with a three-way contrast (/t/, /t'/, and /p/) can be thought of going a step further and neutralizing the /p/ vs. /p'/ distinction altogether.

(2.7)

	cor-V ta vs. t'a	lab-V pa vs. p'a	V-cor# at vs. at'	V-lab# ap vs. ap'
Pattern 1	yes	yes	yes	yes
Pattern 2	yes	yes	yes	
Pattern 3	yes	yes		

These observations can be restated as implicational markedness statements about the environment, palatalization and place contrasts (2.8-2.10). The statements are formulated both in terms of the contrast (2.8a and 2.10a) and in terms of particular consonants (2.8b, 2.9, and 2.10b). The statements below are to be read as follows: a contrast between plain and palatalized consonants before a vowel (or in onset) is less marked than the same contrast in final position (2.8a). Or, the environment before a vowel is less marked for a palatalized consonant than final position.

(2.8) Environment: onset vs. final coda

a. C vs. C'/_# > C vs. C'/_V
• If a language maintains a plain-palatalized contrast in the final coda, it also has it in onset;

b. C'/_# > C'/_V
• If a language has a palatalized consonant in the final coda, it also has it in onset;

(2.9) Consonants: plain vs. palatalized

C'/_# > C/_#
• If a language has a palatalized consonant in the final coda, it also has plain consonants in this position;

(2.10) Consonants: labial vs. coronal

a. C_{lab}/_# vs. C_{lab}/_# > C_{cor}/_# vs. C'$_{cor}$/_#
• If a language maintains a plain-palatalized contrast in labials in the final coda, it also has it in coronals in this position.

b. C'$_{lab}$/_# > C'$_{cor}$/_#
• If a language has a palatalized labial in the final coda, it also has a palatalized coronal in this position.

2.1.3.1.1.3 The direction of sound change

As discussed in section 2.1.2.4, one might expect synchronic asymmetries to be mirrored in phonological change. It is interesting to note that many of the languages that disallow the final /t'/ and/or /p'/ had these segments in this position at earlier historical stages. For instance, Bulgarian (Standard) used to have both /p'/ and /t'/ finally (Carlton 1990: 306). Polish (Carlton 1990: 251-252), Belorussian (Carlton 1990: 299), and Russian dialects went through the process of depalatalization of the final /p'/ (Borkovskii & Kuznetsov 1965: 166, Avanesov & Orlova 1965: 88). Many of the three-way contrast languages (Czech, Slovak, Ukrainian, north-eastern Polish dialects, Scots Gaelic, and Manx) originally had the plain-palatalized distinction in labials. The common process of depalatalization involved a change of the final /p'/ to /p/ and the initial /p'/ to the sequence /pj/ (with possible fronting of the following vowel, as in Czech (Carlton 1990: 234)). Less common cases showed a change of both onset and final coda /p'/ to sequences of [pç], [pɕ], and [pʃ], as in north-eastern Polish dialects (Basara et al., Dejna 1993, Zdunska 1965). Thus, we can say that the general direction of historical change was that labials underwent depalatalization prior to the coronals.

The statements made in (2.8-2.10) can be also applied to the direction of sound change: if palatalized consonants undergo depalatalization, final palatalized labials are likely to be affected first, followed by final palatalized coronals. The opposite process is very unusual.

In sum, sound change in terms of the plain-palatalized contrast appears to proceed gradually in the direction of depalatalization from the least restricted pattern (Pattern 1) to the more restricted ones (Patterns 2 and 3). The palatalized labials are the first to undergo the process of depalatalization, while the palatalized coronals are more resistant to the change.

2.1.3.1.2 Medial environment (onset)

As discussed at the beginning of section 2.1.3, in most of the languages surveyed here the distribution of plain and palatalized consonants did not differ depending on whether the onset is word-initial or medial. A dialect of Karelian (Arkhangelsk: Leskinen 1968), however, presents a case where the /t/ vs. /t'/ contrast occurs medially but not initially. Thus the contrast is neutralized in the word-initial onset, but not in the word-medial onset position. Also, in Lithuanian, /p/ and /p'/ (as [p^j]) are attested medially, while /p/ and the sequence /p/ + /j/ are found initially. In this case, however, the contrast between /p/ and /p'/ is not neutralized completely, but rather transformed into the difference between a segment and a cluster.

Overall, the languages do not treat the two onset positions differently. These two exceptional cases, however, will be important for our discussion.

2.1.3.2 Distribution: Consonants in clusters

The examination of consonants in medial two-consonant clusters is divided into according to the following contexts of occurrence: before plain consonants (coda), before palatalized consonants (coda), and after a consonant (onset). Since not all of the languages examined include possible medial clusters, only some of the languages are discussed.

As before, I use nonsense examples (*apta*, **ap'ta*, etc.) to illustrate well-formed and ill-formed strings. I examine clusters with /p/, /p'/, /t/, and /t'/ as C_1 and C_2, in all possible combinations. I also include clusters with /k/ and /k'/, where available, as C_2, to determine the relevance of the hetero-organic/homorganic distinction. These nonsense words present a certain degree of generalization that would allow for a comparison of the patterns. For instance, /p/, /t/, /p'/, and /t'/ as C_1 can stand for either voiceless or voiced stops (e.g. *tp* may stand for the Russian clusters /tp/ and /db/). Since not all languages allow geminates, the C_2 of *pp* or *t't* can stand for any homorganic obstruent or nasal. Also, the morpheme structure, stress, and vowel quality are not taken into account (but see sections 2.2.2.2.1.4 and 2.2.2.2.2.4 on Russian).

2.1.3.2.1 Before plain consonants (coda)

2.1.3.2.1.1 Data

Sequences with all 12 theoretically possible medial combinations of stops are given in (2.11). In these strings, /p/, /p'/, /t/, and /t'/ occur before plain labials, coronals, and dorsals. The clusters can be hetero-organic (e.g. *p't* and *p'k*) or homorganic (e.g. *p'p*). This distribution pattern (Pattern 1) is reported for an eastern Bulgarian dialect (Nova Nadezhda: Khristov 1956, Chekman 1970): the contrast between plain and palatalized labials and coronals is maintained regardless of the place of articulation of the following consonant.[9] Erzya Mordva, with a three-way contrast (/p/, /t/, /t'/), does not seem to restrict the occurrence of the palatalized /t'/ before plain segments either (Redei 1984: 214-217).

(2.11) Pattern 1: Bulgarian (Nova Nadezhda)

	p	t	k
p	*appa*	*apta*	*apka*
t	*atpa*	*atta*	*atka*
p'	*ap'pa*	*ap'ta*	*ap'ka*
t'	*at'pa*	*at'ta*	*at'ka*

Most of the languages, however, are much more restrictive in the distribution of the palatalized segments. Pattern 2 (2.12) is similar to Pattern 1, but it neutralizes the distinction between preconsonantal labials, where the consonant that surfaces is plain [p]. This pattern is found in other Bulgarian dialects (Krinichnoe: Chekman 1970).

[9] Although the author (Khristov 1956) mentions no restrictions on the palatalized consonants, the examples given do not have palatalized segments before labials. Many (but not all) of the C'C cases involve inflectional morpheme boundaries (Joseph Schallert p.c.).

(2.12) Pattern 2: Bulgarian (Krinichnoe)

	p	t	k
p	*appa*	*apta*	*apka*
t	*atpa*	*atta*	*atka*
p'			
t'	*at'pa*	*at'ta*	*at'ka*

A more common pattern, Pattern 3 (2.13), is attested in Russian, Belorussian, and Polish. In these languages neutralization affects not only labials (to [p]), but also /t'/ before homorganic consonants (e.g. **at'ta*). In this case the plain /t/ is the output of neutralization. Note that the hetero-organic clusters with /t'/ (*t'p* and *t'k*) are perfectly well-formed in these languages. The neutralization pattern is often reflected in alternations /p/ vs. /p'/, /t/ vs. /t'/ in derivational morphology.

(2.13) Pattern 3: Belorussian, Russian (Standard and Cherdyn'), Polish

	p	t	k
p	*appa*	*apta*	*apka*
t	*atpa*	*atta*	*atka*
p'			
t'	*at'pa*		*at'ka*

Finally, many languages, like Irish, Scots Gaelic, Lithuanian, Nenets, and some northern Russian (Vologda and Vyatka regions: Azarkh 1967, 1973) and Bulgarian dialects (Bela Cherkva: Chekman 1970), allow only plain-plain clusters, while prohibiting palatalized-plain sequences (2.14). This pattern, Pattern 4, neutralizes the distinction before plain consonants altogether: only plain segments are possible before other plain consonants.

(2.14) Pattern 4: Irish, Scots Gaelic, Lithuanian, Bulgarian (Standard and Bela Cherkva), Russian (Vologda-Vyatka), Nenets

	p	t	k
p	*appa*	*apta*	*apka*
t	*atpa*	*atta*	*atka*
p'	[redacted]	[redacted]	[redacted]
t'	[redacted]	[redacted]	[redacted]

Some languages with a three-way contrast (/p/, /t/, and /t'/), Ukrainian, Slovak, and Mazovian dialects of Polish, exhibit the same distribution of /t'/ as in Pattern 3. As languages in Pattern 4, Czech (with a few exceptions), Manx, Karelian dialects (Arkhangelsk, Olonetsk 1 and 2) disallow /t'/ before plain consonants. Japanese (a four-way contrast) also presents a sub-case of Pattern 4, showing the same restriction on /p'/ and /t'/ in homorganic clusters (**ap'pa*, **at'ta*, but *appa, atta*). Note that hetero-organic sequences do not occur in this language at all (Vance 1987, Akamatsu 1997).

2.1.3.2.1.2 Summary and generalizations

The overall differences between preconsonantal labials and coronals in neutralization patterns are summarized in (2.15). Pattern 1 allows the contrast in both labials and coronals. Pattern 2 is more restrictive: the labials are neutralized in favour of the plain /p/. Pattern 3 shows the neutralization of labials and a partial restriction on coronals (before homorganic segments). And finally, Pattern 4 imposes a full prohibition against palatalized consonants in the preconsonantal coda.

(2.15)

	cor-C *atpa vs. at'pa*	lab-C *apta vs. ap'ta*
Pattern 1	yes	yes
Pattern 2	yes	[redacted]
Pattern 3	yes / no (t)	[redacted]
Pattern 4	[redacted]	[redacted]

Looking at the neutralization patterns in terms of the place of articulation of the following consonant (2.16), we see that labials are not sensitive to the place of C_2, being either allowed or disallowed before hetero- or homorganic consonants. Coronals, however, tend to maintain the contrast better in the hetero-organic context.

(2.16)

a. Labials

	_Cht *apta vs. ap'ta*	_Chm *appa vs. ap'pa*
Pattern 1	yes	yes
Patterns 2, 3, 4		

b. Coronals

	_Cht *atpa vs. at'pa*	_Chm *atta vs. at'ta*
Patterns 1, 2	yes	yes
Pattern 3	yes	
Pattern 4		

It is important to note that, as in final coda position, in all of the cases before a plain consonant the contrast is neutralized in favour of the plain stop.

These observations can be restated as implicational markedness statements about the environment, palatalization and place contrasts (2.17-2.19). The statements are formulated both in terms of the contrast (2.7a and 2.9b) and in terms of particular consonants (2.7b, 2.8, and 2.9b).

(2.17) Environment: hetero-organic (ht) vs. homorganic (hm)

a. C vs. C'/_Cht > C vs. C'/_Chm
• If a language maintains a plain-palatalized contrast (coronals) before a homorganic segment, it also has it in the hetero-organic context;

b. C/_Cht > C'/_Chm
• If a language has a palatalized segment (coronal) before a homorganic segment, it also has it in the hetero-organic context;

(2.18) Consonants: plain vs. palatalized

C'/_C > C/_C
• If a language has a palatalized consonant before a plain segment, it also has a plain consonant in this environment;

(2.19) Consonants: labial vs. coronal

a. p vs. p'/_C > t vs. t'/_C
• If a language maintains a plain-palatalized contrast before a plain segment in labials, it also has it in this position in coronals.

b. p'/_C > t'/_C
• If a language has a palatalized labial before a plain segment, it also has a palatalized coronal in this position.

2.1.3.2.1.3 The direction of sound change

Recall that in many of the languages in the sample palatalized consonants in preconsonantal position arosc through the syncope of front vowels. This process seems to apply blindly, without reference to the place of articulation of preceding or following consonants (Borkovskii & Kuznetsov 1965: 99-103, Kiparsky 1979: 97-103, Russel 1995). The logical consequence of this would be an emergence of all possible palatalized-plain clusters (as in Pattern 1 (2.11)). Interestingly, none of the languages preserved that state of affairs (in the Bulgarian dialects described in Patterns 1 and 2 the C'C-clusters evolved through a more recent process of syncope (Khristov 1956, Chekman 1970)). Moreover, the depalatalization process targeted the same (or almost the same) sets of clusters. For instance, the East and West Slavic groups of languages came to share Pattern 3 (2.13) independently, since the process of syncope applied after the Common Slavic period (Carlton 1990: 171-172).

In sum, as in the final environment, sound change tends to depalatalize preconsonantal palatalized segments. In this process the palatalized labial is the primary target. /t'/ is more susceptible to depalatalization before hetero-organic consonants than before homorganic ones. Thus, we can expect change to proceed from the least restricted pattern (Pattern 1) to the more restricted ones (Patterns 2, 3, and further 4).

2.1.3.2.2 Before palatalized consonants (coda)

2.1.3.2.2.1 Data

In this section I investigate the distribution patterns found before palatalized consonants. The sequences with all 12 logically possible medial combinations of stops are given in (2.20). In these strings, /p/, /p'/, /t/, and /t'/ occur before palatalized labials, coronals, and dorsals. The clusters can be hetero-organic (e.g. *pt'* and *pk'*) or homorganic (e.g. *pp'*). This distribution pattern with absolutely no restrictions is not attested in any of the examined languages.

(2.20) An unattested pattern

	p'	t'	k'
p	*app'a*	*apt'a*	*apk'a*
t	*atp'a*	*att'a*	*atk'a*
p'	*ap'p'a*	*ap't'a*	*ap'k'a*
t'	*at'p'a*	*at't'a*	*at'k'a*

The least restrictive pattern that is actually found among the languages under consideration is given in (2.21). It is attested in a variety of Standard Russian (Avanesov 1972: 145-168).

(2.21) Pattern 1: Standard Russian

	p'	t'	k'
p		*apt'a*	*apk'a*
t	*atp'a*		*atk'a*
p'	*ap'p'a*		
t'	*at'p'a*	*at't'a*	*at'k'a*

There are two types of restrictions in this pattern. First, only plain labials are permitted before palatalized *hetero-organic* consonants (i.e. **p't'* and **p'k'*). Second, both plain labials and coronals are disallowed before palatalized *homorganic* segments (i.e. **pp'* and **tt'*). Overall, the contrast is maintained only among coronals before hetero-organic consonants (e.g. *atp'a* vs. *at'p'a* and *atk'a* vs. *at'k'a*). It is neutralized before homorganic segments and in all the contexts for labials. Recall that this is exactly the neutralization pattern found in Russian before plain consonants. The difference, however, is in the outcome of neutralization. While it was only plain /p/ or /t/ in the latter case, here both plain and palatalized stops are the outcome, depending on both C_1 and C_2. The details of this pattern, commonly referred to in Russian and Slavic literature as 'secondary regressive palatalization, or softening' (Carlton 1990: 291, Timberlake 1993: 829-830, Hamilton 1980: 68-69), will be examined in more detail in section 2.2.2.2.2.

In Pattern 2 (2.22), found in Irish, Lithuanian, and a Bulgarian dialect (Nova Nadezhda), only palatalized consonants occur before other palatalized segments. Sequences of type /pt'/, /pp'/, /tp'/, or /tt'/ are inadmissible.

(2.22) Pattern 2: Irish, Lithuanian, Bulgarian (Nova Nadezhda)

	p'	t'	k'
p			
t			
p'	*ap'p'a*	*ap't'a*	*ap'k'a*
t'	*at'p'a*	*at't'a*	*at'k'a*

Thus, the plain-palatalized contrast is neutralized in all of the contexts before a palatalized segment. Recall that in Irish and Lithuanian the opposite was attested before plain consonants: only plain-plain clusters were permitted. In other words, all of these languages exhibit a more general restriction: consonants in clusters agree in secondary articulation. It is the quality of the second consonant in the cluster that determines whether the sequence is plain or palatalized, when morphemes are concatenated.

Japanese shows a sub-type of this pattern, with an additional restriction on hetero-organic clusters. Other than that, the restriction on plain-palatalized clusters (as well as palatalized-plain sequences) is identical to that found in Irish and Lithuanian (**app'a*, **att'a*, but *ap'p'a*, *at't'a*).

Some Lithuanian dialects keep the restriction *CC' on coronals, but not on labials (Pattern 3: Ambrazas 1997), as shown in (2.23). Instead, the palatalized labial is prohibited before other palatalized segments.

(2.23) Pattern 3: Lithuanian dialects

	p'	t'	k'
p	*app'a*	*apt'a*	*apk'a*
t	[redacted]	[redacted]	[redacted]
p'	[redacted]	[redacted]	[redacted]
t'	*at'p'a*	*at't'a*	*at'k'a*

Pattern 4 (Standard Bulgarian and dialects) applies the restriction to all of the palatalized segments in this environment (2.24). Only plain consonants are found here as C_1.

(2.24) Pattern 4: Bulgarian (Standard and dialects)

	p'	t'	k'
p	*app'a*	*apt'a*	*apk'a*
t	*atp'a*	*att'a*	*atk'a*
p'	[redacted]	[redacted]	[redacted]
t'	[redacted]	[redacted]	[redacted]

Although the restrictions in Patterns 2, 3, and 4, are applied to different sets of segments, their outcome is the same: the plain-palatalized contrast is completely neutralized before palatalized consonants. The choice of the default segment in this position can be either plain or palatalized, depending on the segment, and on the language.

2.1.3.2.2.2 Summary and generalizations

The overall differences between labials and coronals before palatalized consonants are summarized in (2.25). Pattern 1 allows the contrast in coronals, but neutralizes it in labials. The outcome of neutralization is either /p/ or /p'/ depending on the place of articulation of C_2. Patterns 2, 3, and 4 impose the prohibition on both labials and coronals. Although neutralization may be in favour of either plain or palatalized segments, /t'/ is more common than /t/ as a result of the neutralization of coronals, while /p/ is preferred as the outcome of this process in labials.

(2.25)

	cor-C' *atp'a vs. at'p'a*	lab-C' *apt'a vs. ap't'a*
Patterns 1	yes	[redacted]
Patterns 2, 3, 4	[redacted]	[redacted]

The tables in (2.26) show the neutralization patterns by place of articulation. The place of the following segment plays a certain role in the distribution of coronals and the realization of labials as C_1; however, it is not as important as before plain consonants, and is limited to Pattern 1.

(2.26)
a. Labials

	_Cht *apt'a vs. ap't'a*	_Chm *appa vs. ap'pa*
Pattern 1	[redacted]	[redacted]
Pattern 2	[redacted]	[redacted]
Patterns 3, 4	[redacted]	[redacted]

b. Coronals

	_Cht *atp'a vs. at'p'a*	_Chm *att'a vs. at't'a*
Pattern 1	yes	[redacted]
Patterns 2, 3	[redacted]	[redacted]
Pattern 4	[redacted]	[redacted]

These observations can be restated as implicational markedness statements in (2.27-2.29).

(2.27) Environment: hetero-organic vs. homorganic (coronal)

a. C vs. $C'/\_C'_{hm}$ > C vs. $C'/\_C'_{ht}$
• If a language maintains a plain-palatalized contrast before a homorganic palatalized segment, it also has it in the hetero-organic context;

b. $C'/\_C'_{hm} > C/\_C'_{ht}$
• If a language has a plain segment before a homorganic palatalized segment, it also has it in the hetero-organic context;

(2.28) Consonants: labial vs. coronal

a. p vs. p'/_C' > t vs. t'/_C'
• If a language maintains a plain-palatalized contrast before a palatalized segment in labials, it also has it in this position in coronals.

b. p'/_C' > t'/_C'
• If a language has a palatalized labial before a palatalized segment, it also has a palatalized coronal in this position.

(2.29) Consonants: plain vs. palatalized (Patterns 1-3 with some limitations)

C'/_C > C/_C
• If a language has a plain consonant before a palatalized segment, it also has a palatalized consonant in this environment;

Overall, the environment before palatalized consonants is more restrictive than that before plain segments. Neutralization can result in either palatalized or plain segments. The palatalized option is more likely for coronals than labials. Two consonants are more likely to agree with respect to palatalization if they are of the same place of articulation.

2.1.3.2.2.3 The direction of sound change

Given the fact that CC' and C'C clusters arose through the process of syncope, we would expect the languages to exhibit all possible clusters of these types. As we saw, however, the range of possible contrasts is even smaller than in the previously examined case. The languages, however, show a wide range of variation with regard to which clusters are permitted. Within a language this environment appears to be the least stable for palatalization. For example, the prescribed norm in Russian with respect to the pronunciation of some CC' and C'C' clusters has changed several times since the beginning of the 20th century (Avanesov 1972: 145-146, 240-242).

In sum, changes in consonants before palatalized segments show less consistency: they can be either palatalized, in agreement with the following consonant, or depalatalized. The first process seems to prevail. One way or another, the change tends to lead to the elimination of contrast in this environment.

2.1.3.2.3 After consonants (onset)

Our examination of the two preconsonantal positions showed that the first consonant in a cluster is subject to a number of restrictions. Its realization, to a large extent, depends on the quality (mostly with respect to palatalization, but also place) of the following segment. Does the same ever hold for the second consonant in the cluster? Is it dependent on whether C_1 is plain or palatalized, labial or coronal? This can be determined only by examining the neutralization patterns in derived environments. It appears to be the case, however, that while regressive assimilation of palatalization is quite common in these languages, the progressive process is rather rare.[10] Overall, the second consonant of a two-segment cluster patterns with respect to palatalization the same way as single onset consonants (initial and medial). Thus, everything said about the onset position in section 2.1.3.1.1 holds for C_2.

2.1.3.3 Distribution: Summary

In the previous sections we examined the distribution of plain and palatalized labial and coronal stops in syllable onset and three syllable coda environments: final position, before plain consonants and before palatalized consonants. In (2.30) I present the overall results showing the range of attested types with respect to neutralization of the contrast. The types are given in the left column of (2.30a) and the environments are shown at the top. Type 1 shows no neutralization

[10] Some Russian dialects exhibit progressive palatalization of the velar /k/ after a palatalized consonant. This case is largely lexically restricted to the diminutive suffix -k- (Avanesov & Orlova 1965: 74-75, Borkovskii & Kuznetsov 1965: 113-115).

("yes"), allowing the contrast in all the environments. Type 2 imposes a restriction on the context before palatalized segments ("no"), but not in other positions. Type 3 disallows the contrast before consonants, but permits it finally and before a vowel. And finally, Type 4 limits the contrast to the onset position only and neutralizes it in all the other contexts.

Tables 2.30bc list the languages that correspond to these types separately for labials (b) and for coronals (c). Note that all of the types listed are attested, except for Type 1 in labials. Type 1 in coronals, however, is not without additional restrictions. Recall that the languages listed under this type maintain the preconsonantal contrast only in the hetero-organic position. In this respect, the languages listed under Type 2 (coronal) show a wider range of distribution before plain consonants.

Interestingly, Type 4 (onset only) for labials is shared by most of the languages, while coronals show more variability.

(2.30)

a.

	_V *pa vs. p'a*	_# *ap vs. ap'*	_C *apta vs. ap'ta*	_C' *apt'a vs. ap't'a*
Type 1	yes	yes	yes	yes
Type 2	yes	yes	yes	
Type 3	yes	yes		
Type 4	yes			

b. Labials

Type 1	*unattested*
Type 2	Bulgarian (NN)
Type 3	Irish, Russian
Type 4	Belorussian, Bulgarian, Bulgarian (K), Japanese, Lithuanian, Nenets, Polish, Russian (Ch), Russian (VV), Scots Gaelic
No contrast	Czech, Karelian (A), Karelian (O1), Karelian (O2), Polish (M), Slovak, Ukrainian

c. Coronals

Type 1	Belorussian, Polish, Russian, Ukrainian
Type 2	Bulgarian (NN), Bulgarian (K), Erzya Mordva
Type 3	Irish, Scots Gaelic, Manx, Czech, Slovak, Bulgarian (BCh), Karelian (O1)
Type 4	Bulgarian, Lithuanian, Nenets, Russian (VV), Karelian (O2), Karelian (A), Japanese

Note: --Bulgarian (NN): Nova Nadezhda dialect; Bulgarian (K): Krinichnoe dialect; Bulgarian (BCh): Bela Cherkva dialect;
--Karelain (A): Arkhangelsk dialect; Karelain (O1 and O2): Olonetsk dialects; Karelain (A) has the contrast in medial onset only; Polish (M): Mazovian dialects of Polish;
--Russian (Ch): Cherdyn' dialect, Perm' region; Russian (VV): dialects of Vologda and Vyatka regions;
--In Japanese the stops do not occur in the final environment.

The type a language belongs to with respect to the contrast in labials is either more restricted in labials than in coronals, or identical. Notice the gap with no cases with more restricted coronals. This is summarized in (2.31). "Labial > coronal" (a) means that the contrast /p/ vs. /p'/ in certain environments implies the distinction /t/ vs. /t'/ in the same positions. Or, the latter contrast occurs in the same or in a wider set of contexts than the former one. The statement in (b) stands for no preference for either contrast, and the one in (c) shows the wider distribution of plain-palatalized labials. None of the languages shows a distributional preference for labials at the expense of coronals (c).

(2.31) a. labial > coronal

Belorussian, Russian, Polish, Scots Gaelic, Russian (NW), Bulgarian (K)

b. labial, coronal
Bulgarian (NN), Bulgarian, Irish, Nenets, Lithuanian, Russian (V), Japanese

c. coronal > labial
unattested

Recall that the languages without palatalized labials (2.32) can also be treated as a sub-case of (2.31), i.e. showing a preference for coronals.

(2.32) labial > coronal (no palatalized labials)
Czech, Erzya Mordva, Karelian, Slovak, Ukrainian

Based on all our observations in this section we can make the following implicational statements about the distribution and neutralization of the plain-palatalized contrast in labials and coronals (summarized from various sections).

(2.33) Environment

a. C vs. C'/_C' > C vs. C'/_C > C vs. C'/_# > C vs. C'/_V
• If a language maintains a plain-palatalized contrast lower on the scale, it also does so in all other environments higher on the scale

b. $C'/\_C_{hm} > C/\_C_{ht}$ > or $C'/\_C'_{hm} > C/\_C'_{ht}$ (coronals)
• If a language has a palatalized segment before a homorganic segment (C or C'), it also has it in the hetero-organic context;

(2.34) Consonants: labial vs. coronal

a. p vs. $p'/\_]_{\sigma}$ > t vs. $t'/\_]_{\sigma}$
• If a language maintains a plain-palatalized contrast in labials in any of the coda environments, it also has it in coronals in these positions.

b. $p'/\_]_{\sigma} > t'/\_]_{\sigma}$
• If a language has a palatalized labial in any of the coda environments, it also has a palatalized coronal in these positions.

(2.35) Consonants: plain vs. palatalized

a. C'/_#, C > C/_#, C
• If a language has palatalized consonants in the final coda or before a plain consonant, it also has plain consonants in these positions;
• If a language has palatalized consonants in coda, it also has plain consonants in this position;

b. C/_C' > C'/_C' (more likely for coronals and homorganic segments)
• If a language has a consonant before a palatalized consonant, it is likely to be also palatalized;

In sum, the examination of distribution patterns for the consonants /p/, /p'/, /t/, and /t'/ in a number of environments shows that there are a number of asymmetries with respect to position, place, and palatalization. The onset environment is the least restricted, while the context before a palatalized consonant is the most prohibitive. Palatalized coronal /t'/ occurs in more environments than palatalized labial /p'/. The final and preconsonantal coda (_C) positions show a preference for a plain consonant, while the environment before a palatalized segment may accept either palatalized (preferred) or plain consonants.

2.1.4 Conclusion

We have seen in this section that there are important differences between palatalized labials and coronals. The two consonants differ in the terms of their occurrence in language inventories, the range of phonetic realization, and in their distribution patterns. We have also seen that there are some environments that are 'good' for the plain-palatalized contrast and some that are 'bad'. Which consonant is likely to surface in a particular environment depends on the place and secondary articulation of this consonant as well as the context.

In the next section I examine some of the patterns described above in more detail, illustrating them with data from Russian.

2.2 Phonotactics of Russian contrastive palatalization

The major topic of this section is an examination of the foundations of the plain-palatalized contrast in Standard Russian. Before turning to the phonetic basis for this contrast I examine in greater depth the categorical restrictions, asymmetries in phoneme frequency, and morphological conditioning of the contrast.

The data come from a number of sources (Townsend 1968/1975, Tolstaia 1968, Jones & Ward 1969, Hamilton 1970, Avanesov 1972, Zalizniak 1977, Kiparsky 1979, Dmitrienko 1985), and are partly supplemented by the author. The frequency counts (overall, initial, final, and in stop clusters) are based on a partly transcribed (consonants only) corpus of written Russian (Kochetov, in preparation) that comprises about 33,000 words. I also used the reverse Russian dictionary by Zalizniak 1977 to determine the number and morpheme structure of lexical items with the stops of interest.

I show in this section that the frequency and morphological information complement the categorical restrictions discussed in section 2.1, providing insights into the asymmetries between plain and palatalized consonants, labials and coronals, and the relevant environments.

2.2.1 Plain-palatalized contrast

The Russian consonant inventory is given in (2.36). The language can be considered as representative of the full possibilities of palatalization. The plain-palatalized phonemic distinction involves all places of articulation: labials, coronals and velars. The consonants of various manners of articulation (stops, affricates, fricatives, nasals, and liquids) can also be either plain or palatalized. Most of the segments are paired with respect to palatalization, while a small number are phonetically either plain (/ts/, /ʂ/, /ʐ/, and /x/) or palatalized (/tʃ/, /ʃ:/, and /j/), without a counterpart. Phonologically these also tend to pattern with either plain or palatalized segments.

(2.36) Russian consonant phoneme inventory

[illegible]		[illegible]				[illegible]	
plain	pal	plain	pal	plain	pal	plain	pal
p	**p'**	**t**	**t'**			k	k'
b	b'	d	d'			g	g'
		ts			tʃ		
f	f'	s	s'	ʂ	ʃ:	x	[x']
v	v'	z	z'	ʐ			
m	m'	n	n'				
		l	l'				
		r	r'				
					j		

As in the previous section, I limit the discussion to the distribution of coronal and labial stops (given in bold in (2.1)), disregarding laryngeal distinctions.

2.2.1.1 Phonetic realization

Details of the articulation of Russian plain and palatalized consonants are discussed in Chapter 5. In this section I limit the discussion by saying that the Russian plain /p/ is realized either as a plain or velarized labial, while /p'/ is characterized by a secondary palatal articulation

(2.37). The plain coronal /t/ is described as apico-laminal dental, possibly velarized. Its palatalized counterpart has a laminal constriction at the alveolar region (e.g. Bondarko 1981).

(2.37) Realizations of plain and palatalized voiceless labial and coronal stops

/p/	/p'/	/t/	/t'/
[p] / [pˠ]	[pʲ]	[t̪]/[t̪ˠ]	[tʲ]

It is important to note that Russian has fairly typical phonetic realizations of palatalized stops that differ from the plain counterparts primarily in the presence of a secondary palatal articulation.

2.2.1.2 Phoneme frequency

The relative frequency of plain and palatalized labial and coronal stops is shown in (2.38). The percentage is based on the overall occurrence of the four stops (23,726 counts: /p/ (6,152), /p'/ (1,646), /t/ (10,937), /t'/ (4,991)). Since the voice distinction is not important for our purposes (and is neutralized finally), the numbers for both voiceless and voiced stops are combined.

(2.38) Relative frequency of plain and palatalized labial and coronal stops (voiced and voiceless combined) in the corpus (all positions)

/p/	/p'/	/t/	/t'/
26%	7%	46%	21%

We can see that the plain coronal /t/ is the most frequent of the four stops, accounting for almost half of the counts (46%). The palatalized labial /p'/ is the least frequent segment (7%). For both labials and coronals, plain consonants are more common than their palatalized counterparts. In other words, the frequency of plain and palatalized consonants in Russian correlates with the overall restrictions on these segments cross-linguistically. The same can be said about the difference between palatalized /p'/ and /t'/. As we will see later, even in the 'best' environment for palatalization, onset position, /p'/ is the least frequent segment.

2.2.2 Distribution

In the sections below I examine the distribution patterns of Russian /p/, /p'/, /t/, and /t'/ as single consonants and as segments in clusters. Again, the questions of interest are: what stops are more susceptible to the neutralization of the plain-palatalized contrast, what is the outcome of this neutralization, and what environments induce this process?

2.2.2.1 Single consonants

2.2.2.1.1 Pattern

In this section I consider the distribution patterns of plain and palatalized labials and coronals as single consonants.

Recall that in terms of the distribution of single consonants Russian exhibits Pattern 1 (2.39), allowing all four consonants in both initial and final coda positions.

(2.39) Pattern 1 (No restrictions): Russian (repeated from (2.4)

	p	t	p'	t'
onset	*pa*	*ta*	*p'a*	*t'a*
final coda	*ap*	*at*	*ap'*	*at'*

Below I examine the distribution of labials and coronals separately.

2.2.2.1.2 Labials

Initial and medial plain and palatalized labials are illustrated in (2.40).[11] The consonants occur in all vocalic environments, before and after any vowel, although there are certain lexical limitations on the distribution of plain consonants before front vowels, where palatalized consonants are found more frequently (Avanesov 1972: 139-140).

[11] Here and below I use the transliteration adopted in the North American literature on Russian, while using the IPA symbols for transcription. Note the following correspondences: ë = [o] (C'__) , y = [i].

(2.40) a. initial

/p/		/p'/	
[p]at	*stalemate*	[p^j]at	*heel*, gen., pl.
[p]ot	*sweat*	[p^j]ëk	*baked*
[p]ugat'	*to scare*	[p^j]ure	*puree*
[p]er	*peer*	[p^j]erl	*pearl*
[p]yl'	*dust*	[p^j]il	*drank*

b. medial

/p/		/p'/	
ko[p]at'	*dig*	o[p^j]at'	*again*
sa[p]og	*high boot*	sa[p^j]ër	*combat engineer*
lo[p]ux	*burdock*	ku[p^j]ura	*banknote*
kana[p]e	*canapé*	te[p^j]er'	*now*
u[p]yr'	*vampire*	u[p^j]irat'	*to prop*

It should be noted that the palatalized stops (both labials and coronals) occur more often in medial onset position, especially before inflectional and derivational suffix boundaries beginning with vowels.

Labials in the final coda position are shown in (2.41). Again, the quality of the preceding vowel does not affect the distribution. Palatalization has a certain functional load in this environment: the addition of the secondary articulation to a stem (e.g. *grab-*, 'to rob') derives an imperative verb form (*grab'*). Most of the other items are feminine nouns (e.g. *syp'* 'rash' and *cep'* 'chain'). It should be mentioned that the final environment neutralizes the voice distinction in Russian. The underlying consonants are contrastive prevocalically in inflectional and derivational affixes.

(2.41) final

/p/		/p'/	
gri[p]	*mushroom*	progi[p^j]	*caving in*
ry[p]	*fish*, gen, pl	sy[p^j]	*rash*
ce[p]	*flail*	ce[p^j]	*chain*
gra[p]	*hornbeam*	gra[p^j]	*rob*, imp.
glu[p]	*stupid*	glu[p^j]	*depth*
uto[p]	*has sunk*	to[p^j]	*swamp*

Russian also contrasts palatalized /p'/ with sequences of /p/ or /p'/ plus yod, as shown in (2.42). /p'/ co-occurs with /p'/+/j/ (as [p^{j}j]) stem-internally before vowels (2.42a). The sequence of plain /b/ with /j/ (as [bj] or [b^{j}j], depending on the variety of Standard Russian) is found across prefix-stem boundaries (Avanesov 1972: 140, 166-167). In final coda position both /p'/ and the sequence /jp/ are distinguished (2.42c). However, there is no final cluster of /j/ with palatalized labial /p'/.

(2.42)	/p'/		/p'j/, /pj/, or /jp/	
a.	[p^j]atyj	*fifth*	[p^{j}j]anyj	*drunk*
	sa[p^j]ër	*combat engineer*	ko[p^{j}j]ë	*spear*
b.	re[b^j]ata	*kids*	o[bj]atyj	*embraced*
c.	gra[p^j]	*rob*, imp.	sha[jp]	*puck* gen.pl.

Although well-established in both onset and coda positions, /p'/ is relatively rare. It accounts for only 8% of the total occurrence of the four stops in word-initial position (out of 8,292 tokens; based on the corpus mentioned above), while the plain /p/ is found in half of the total number.[12] Word-finally /p'/ is even more rare (0.2%, 8 total; out of 3,524 total); /p/ in this position is not very frequent either (5%).

Based on Zalizniak 1977, the final /p'/ (*-p'* or *-b'*) is found in only 40 lexical items, mostly nouns (mainly feminine) and imperative forms of verbs. In other morphological forms of these nouns and verbs /p'/ occurs before vowels (affixes -i, -a, -e, etc.).

2.2.2.1.3 Coronals

Initial and medial plain and palatalized coronals are illustrated in (2.43). The consonants occur in all vocalic environments, before and after any vowel, with certain lexical limitations before front vowels, where palatalized consonants are preferred (Avanesov 1972: 139-140).

[12] Medial consonants were not counted.

(2.43) a. initial

/t/		/t'/	
[t]ak	*so*	[t^j]aga	*drought*
[t]omnyj	*obese*	[t^j]omnyj	*dark*
[t]ura	*rook*	[t^j]urja	*tjurja* (soup)
[t]est	*test*	[t^j]esto	*dough*
[t]ykat'	*to poke*	[t^j]ikat'	*to tick*

b. medial

/t/		/t'/	
va[t]a	*cotton*	ba[t^j]a	*dad*
po[t]ok	*stream*	po[t^j]ok	*began to flow*
pe[t]ux	*rooster*	u[t^j]ug	*iron*
pa[t]ent	*patent*	po[t^j]erja	*loss*
mo[t]yga	*hoe*	mo[t^j]iv	*tune*

The coronals in final coda position are shown in (2.44). They can be found after any vowel. Both segments in this position play an important morphological role marking either 3rd person present and past passive participle (/t/) or infinitive (/t'/) verb forms (e.g. *pojet* s/he sings, *pet* sung, *pet'* to sing).

(2.44) final

/t/		/t'/	
m[at]	*mat*	m[at^j]	*mother*
d[ut]	*blown*	d[ut^j]	*to blow*
x[ot]	*walk*	x[ot^j]	*though*
b[it]	*beaten*	b[it^j]	*to beat*
p[et]	*sung*	p[et^j]	*to sing*

Russian also differentiates between palatalized /t'/ and the sequence of /t'/ with yod (2.45). The latter is found only stem-internally (2.45a), while the sequence of plain /t/ with /j/ (always as [tj]) is found across prefix-stem boundaries (2.45b) (Avanesov 1972: 166-168). In final coda position both /t'/ and the sequence /jt/ are distinguished (2.45c). However, there is no final cluster of /j/ with palatalized labial /t'/.

(2.45) /t'/		/t'j/ or /tj/	
a. xo[tʲ]a	*although*	sta[tʲj]a	*article*
b. o[tʲ]ec	*father*	o[tj]ezd	*departure*
c. ple[tʲ]	*whip*	fle[jt]	*flute* gen.

Both /t/ and /t'/ have relatively high frequency, which is partly due to their function as markers of verb forms in final position. In each case, however, the plain stop is more frequent than the palatalized one. Thus, /t/ accounts for 33% of the total occurrence of the four stops in word-initial position (out of 8,292 tokens), while palatalized /t'/ is found 10% of the time. Word-finally the figures are 59% for /t/ and 36% for /t'/ (out of 3,524 tokens). Also, the palatalized coronal is more frequent than the palatalized labial (which is also true for the plain stops).

Final /t'/ is also quite frequent in the lexicon. Zalizniak 1977 lists 109 lexical items with final /t'/ (*-t'* or *-d'*) (the infinitive suffix is counted once), mostly feminine nouns. In oblique forms of these words /t'/ occurs before vowels (affixes -i, -a, etc.).

2.2.2.1.4 Summary and conclusion

Morphological conditioning of the occurrence of palatalized consonants is of particular interest. First, these segments in final position often function as markers of certain morphological classes (e.g. gender or declension). Second, final palatalized consonants occur mainly in forms that exhibit alternations where the final palatalized segment can occur before a vowel. Interestingly, palatalized labials in some non-alternating forms were historically depalatalized (Borkovskii & Kuznetsov 1965: 120).

Data from Russian dialects (Avanesov & Orlova 1965: 88, Avanesov & Bromlei 1986: 175, map 70) show that the final coda distinction between /p/ and /p'/ is very commonly neutralized in most Russian dialects, with the exception of the central ones. The same process very rarely affects the coronal stops. The case documented in Azarkh 1967, 1973 for a northern Russian dialect in Vologda and Vyatka regions, where both labials and coronals are depalatalized finally, is rather exceptional. Interestingly, in this dialect the plain stops are realized as apical alveolars (Kuznetsova 1969: 44-60, Azarkh 1973: 144) rather than apico-laminal dentals (as in Standard Russian).

In sum, the analysis of the distribution of Russian single plain and palatalized labial and coronal stops shows that although all these segments are phonotactically unrestricted in both onset and final coda positions, they show frequency asymmetries in terms of palatalization, place and environment, similar to the categorical restrictions found in other languages. The occurrence of palatalized consonants is also found to be partly limited to certain morphological contexts.

2.2.2.2 Consonants in clusters

As before, I examine the distribution of stops before plain consonants (coda), followed by the environments before palatalized consonants (coda) and after a consonant (onset).

2.2.2.2.1 Before plain consonants (coda)

2.2.2.2.1.1 Pattern

Recall that the distribution of four stops in this position in Russian is characterized by Pattern 3, repeated in (2.46). This pattern neutralizes labials in general, suspending the contrast in coronals only before homorganic consonants.

(2.46) Pattern 3: Russian (repeated from (2.13))

	p	t	k
p	*appa*	*apta*	*apka*
t	*atpa*	*atta*	*atka*
p'			
t'	*at'pa*		*at'ka*

2.2.2.2.1.2 Labials

The distribution of preconsonantal (_C) labials in Russian is illustrated in (2.47). The examples have all possible clusters with /p/ (morpheme-internally and across morpheme boundaries). As we see, the palatalized labial /p'/ is not found in any of these contexts.

(2.47)	/p/		/p'/
	le[pt]a	*mite*	*-[p^jt]-
	o[pt]ochit'	*to sharpen*	*-[p^jt]-
	to[pk]a	*furnace*	*-[p^jk]-
	o[pk]osit'	*to mow off*	*-[p^jk]-
	gru[pp]a	*group*	*-[p^jp]-

The phonotactic restriction on /p'/ can be observed in both non-derived (2.47) and derived environments. (2.48) illustrates alternations between /p'/ and /p/: the addition of a suffix (a) or a change in case for some nouns from nominative to plural or oblique forms (b) creates an ill-formed sequence of /p'/ with a plain consonant. This sequence is repaired by neutralizing the contrast in favour of the plain /p/. The place of articulation of the following consonant does not play any role.

(2.48)	a. stem		stem + suffix	
	golu[p^j]	*pigeon*	golu[pk]a	*female pigeon*
	ce[p^j]	*chain*	ce[pn]oj	*chain*, adj.
	b. nominative case		oblique case	
	ku[p^j]ec	*merchant*	ku[p]ca	*merchant*, nom.pl.
	bo[b^j]er	*beaver*	bo[b]ra	*beaver*, gen.sg.

Clusters with /p'/, however, are common across a word boundary, as shown in (2.49), with no restrictions reported (Jones & Ward 1969, Avanesov 1972).

(2.49) word boundary

/pt/	/p't/
gra[p#t]ovarishcha	gra[pʲ#t]ovarishcha
comrade's hornbeam	*rob (imp.) the comrade*
gra[p#k]onduktora	gra[pʲ#k]onduktora
train attendant's hornbeam	*rob (imp.) the train attendant*
gra[p#p]omeshchika	gra[pʲ#p]omeshchika
landowner's hornbeam	*rob (imp.) the landowner*

Word-internal palatalized labials before plain consonants are not found in any Russian dialects (Avanesov & Orlova 1965). There does not seem to be any historical evidence that clusters the type /p'C/ were possible in Old Russian soon after the process of syncope (dropping of *jers*: Kiparsky 1979: 97-103, 126-129).

To summarize, the distinction between plain and palatalized labial stops preconsonantally in words is neutralized altogether within a word, but is allowed across a word boundary. The outcome of neutralization is plain /p/.

2.2.2.2.1.3 Coronals

Recall that the distribution of the plain-palatalized contrast in coronals in Russian is sensitive to the place of articulation of the following consonant. Below I examine the occurrence of the two coronal segments by context: before velars, before labials, and before coronals.

The contrast before a following velar (particularly /k/) is quite common (2.50).

(2.50) preconsonantal: before velars

/t/		/t'/	
ka[tk]a	*pail*	Ka[tʲk]a	*Kate*, familiar
lopa[tk]a	*paddle*	ba[tʲk]a	*dad*, familiar
vzja[tk]a	*bribe*	dja[tʲk]a	*uncle*, familiar
re[tk]o	*rare*	re[tʲk]a	*radish*

These nouns belong to the 1st declension. The pattern does not hold for nouns of the 3rd declension: adding the suffix *-k* leads to depalatalization of /t'/ (2.51). This fact will be of importance in our further analysis.

(2.51)	stem		stem + suffix	
	ni[t^j]	*thread*	ni[tk]a	*thread*, dimin.
	ma[t^j]	*mother*	ma[tk]a	*womb*
	zhu[t^j]	*horror*	zhu[tk]ij	*horrible*

The palatalized /t'/ (and its voiced counterpart /d'/) is attested before labials (2.52). The sequences [d^{j}b], [t^{j}m], [d^{j}m], are contrastive with the sequences [db], [tm], [dm]; however, the latter three occur only across prefix-stem boundaries.

(2.52)	preconsonantal: before labials			
	/t/		/t'/	
	po[db]adrivat'	*to cheer up*	sva[d^{j}b]a	*wedding*
	o[tm]yt'	*to wash off*	po[t^{j}m]a	*dusk*
	po[dm]oga	*help*	ve[d^{j}m]a	*witch*
	po[dm]ostki	*scaffold*	se[d^{j}m]oj	*seventh*

The contrast between /t/ and /t'/ is completely neutralized before coronals (e.g. before /n/ and /s/): only plain /t/ is allowed before these consonants. When words with a final palatalized /t'/ combine with derivational affixes (e.g. high frequency suffixes *-n-*, *-sk-*, and *-stv-*) or other stems beginning with a coronal, the underlying palatalized /t'/ surfaces as a plain /t/ (2.53a). There are no exceptions to this constraint. Plain /t/ does not undergo any changes (2.53b).

(2.53)	stem		stem + suffix	
a.	pu[t^j]	*way*	pu[tn]yj	*appropriate*
	pja[t^j]	*five*	pja[tn]adcat'	*fifteen*
	plo[t^j]	*flesh*	plo[ts]kij	*carnal*
	pja[t^j]	*five*	pja[ts]ot	*fifty*
	my[t^j]	*to wash*	my[ts]a	*to wash oneself*
b.	po[t]	*sweat*	po[tn]yj	*sweaty*
	a[t]	*hell*	a[ts]kij	*hellish*

To summarize, while the plain /t/ is not sensitive to the following consonant, occurring before a large number of consonants, the palatalized /t'/ is restricted. It is allowed only before labials and velars and disallowed before coronals.

Although /t'/ before hetero-organic consonants is perfectly legal, its occurrence is very infrequent compared to its plain counterpart. While the cluster /tp/ is found in the corpus in 90%

(28 tokens) of the total occurrence of all stops in this context (_p), /t'p/ accounts only for 6% (2) of the cases (other cases involve the prefix-stem boundary cluster /tt/). The cluster /tk/ with the plain coronal is very common in the corpus, 74% (150), while /t'k/ occurs only 1% of the time (2) in the appropriate environment (other cases include /pk/). Thus, the contrast between /t/ and /t'/ before plain consonants is very limited and its distribution is always in favour of the plain stop.

The analysis of words with the clusters /t'k/ and /t'p/ (*-d'b-*) (based mainly on Zalizniak 1977 and complemented by the author) shows that most of the 27 lexical items (17 for /t'k/ and 10 for /t'p/) are diminutives, nicknames, and words referring to relatives derived from nouns of the 2nd declension by the suffix *-k-* (e.g. *Ka*[tʲ]*-a* --> *Ka*[tʲ]*-k-a*).[13] All of them have very transparent morpheme boundaries and are often found without suffixes (except for the word *red'ka* 'radish'). In addition, in most of the nouns /t'/ surfaces before a vowel in oblique forms (*jer* (/e/) alternations) or with other derivational suffixes (Townsend 1968/1975). As I noted above, the contrast between /t/ and /t'/ before labial stops is absent in the same morphological environments: while /t'p/ (*-d'b-*) is found in stems, /tp/ (*-tp-* and *-db-*) occurs only in prefix-stem clusters. The contrast before velars is restricted in nouns to one morphological context (derived from nouns of the 1st declension).

All the above-mentioned restrictions apply only within phonological words. Nothing prohibits /t'/ before any consonant of a following word (2.54).[14] Recall that both plain and palatalized labials are also found in this context. This fact will later allow us to examine the full range of clusters in phonetic experiments.

[13] There are also 4 words with the morpheme-internal cluster /t'm/ (*-t'm-* and *-d'm-*).

[14] Or a particle (e.g. *mat'-to* 'mother' (emph.))

(2.54) word boundary

/t/	/t'/
ma[t#n]achal'nika	ma[t'#n]achal'nika
boss's foul language	*boss's mother*
ma[t#s]otrudnika	ma[t'#s]otrudnika
colleague's foul language	*colleague's mother*
ma[t#k]onduktora	ma[t'#k]onduktora
conductor's foul language	*conductor's mother*

Russian dialects do not differ much from Standard Russian in terms of the contrast between /t/ and /t'/ before plain consonants (Avanesov & Orlova 1965, Avanesov & Bromlei 1986). Only the above-mentioned northern dialect (Azarkh 1967, 1973) shows neutralization before all consonants (and in final position).

2.2.2.2.1.4 Summary: Morphological factors

We have seen that the maintenance of the plain-palatalized contrast and, particularly, the occurrence of palatalized consonants is sensitive to certain morphological and lexical factors.

First, the palatalized segments before plain consonants (C'C) are less likely to be found within a morpheme than across morpheme boundaries (stem + suffix), favouring the more morphologically transparent environment. In this respect the word boundary is the best environment for sustaining the contrast. Second, this contrast tends to be restricted to certain lexical classes (e.g. names). Third, the contrast may be suspended in some morphological categories (e.g. 3rd declension) even if the phonological environment is the same as in some class where the contrast is allowed. Fourth (and finally), palatalized consonants are more likely to be found in forms that exhibit alternations where the palatalized C_1 can occur before a vowel.

These seemingly unrelated factors are of importance for the analysis presented in Chapter 8.

2.2.2.2.1.5 Conclusion

To summarize, the distinction between the plain and palatalized coronal stops is maintained before some consonants (e.g. labials and velars) and neutralized before others (e.g. coronals). In the latter context we find only plain stops. The maintenance of contrast before velars and labials is often limited in frequency and is restricted to certain morphological or lexical classes. Only a few clusters that are contrastive with respect to palatalization are morpheme-internal.

2.2.2.2.2 Before palatalized consonants (coda)

2.2.2.2.2.1 Pattern

Recall that with respect to the context before palatalized consonants Standard Russian exhibits Pattern 1 (2.55), allowing the contrast only among coronals before homorganic consonants. Before homorganic segments both labials and coronals surface as palatalized. Labials before hetero-organic consonants are plain.

(2.55) Pattern 1: Standard Russian (repeated from (2.21))

	p'	t'	k'
p	[redacted]	*apt'a*	*apk'a*
t	*atp'a*	[redacted]	*atk'a*
p'	*ap'p'a*	[redacted]	
t'	*at'p'a*	*at't'a*	*at'k'a*

2.2.2.2.2.2 Labials

Plain and palatalized labials before palatalized consonants are illustrated in (2.56). The plain labial /p/ is found before hetero-organic consonants and the palatalized /p'/ occurs before homorganic segments.[15]

(2.56)

a. a[pt^j]eka *pharmacy* *-[p^jt^j]-
ta[pk^j]i *slippers*, loc., sg. *-[p^jk^j]-

b. *-[pp^j]- gru[p^jp^j]e *group*, dat. sg.
*-[bm^j]- o[b^jm^j]en *exchange*

---

[15] The cluster /p't'/ is found with the particle *-te* (imperative, pl) (e.g. *prisposo*[p^jt^j]*e* 'adapt'). In colloquial speech /p'/ in this form is often depalatalized (Avanesov 1972: 138-39).

This restriction can often be observed at the end of a stem when an inflectional affix (usually starting with a front vowel) is added (e.g. *gru*[pp]*a*, 'group' nom. sg., but *gru*[p^jp^j]*e* 'group', dat. sg).

Clusters with /p'/, however, are common across word boundaries, as shown in (2.57), with no restrictions reported (Avanesov 1972, Jones & Ward 1969).

(2.57) word boundary

	/p/	/p'/
	gra[p#p^j]etra	gra[p^j#p^j]etra
	Pjotr's hornbeam	*rob Pjotr*
	gra[p#n^j]ikolaja	gra[p^j#n^j]ikolaja
	Nikolaj's hornbeam	*rob Nikolaj*
	gra[p#k^j]irilla	gra[p^j#k^j]irilla
	Kirill's hornbeam	*rob Kirill*

2.2.2.2.2.3 Coronals

Plain and palatalized coronals before palatalized consonants are illustrated in (2.58). The plain coronal /t/ is found only before hetero-organic consonants, while the palatalized /t'/ occurs before hetero-organic and homorganic segments.

(2.58)

/t/		/t'/	
re[tk^j]ij	*rare*	re[t^jk^j]e	*radish*, prep.sg.
o[db^j]elivat'	*to whitewash*	sva[d^jb^j]e	*wedding*, prep.sg.
*[tt^j]		o[t^jt^j]agivat'	*to postpone*

These clusters are often found at the end of a stem when an inflectional affix (usually starting with a front vowel) is added, or at prefix-stem and preposition-stem boundaries.

Although /t'/ before hetero-organic consonants is perfectly legal, its occurrence is very infrequent compared to its plain counterpart. While the cluster /tp'/ is found in the corpus in 89% (24 tokens) of the total occurrence of all stops in this context (_p'), /t'p'/, although possible, does not occur in the corpus (0%; 0 tokens) (other cases involve the prefix-stem boundary cluster /t't'/). The situation is the same before the velar /k/: the cluster /tk'/ with the plain coronal is relatively common (65%; 35 tokens), while /t'k'/ is not found in the corpus (0%, 0 tokens) (other cases include the quite frequent cluster /pk'/).

The analysis of words with the clusters /t'k'/ and /t'p'/ (*-d'b'-*) (based mainly on Zalizniak 1977) shows that all of the 25 lexical items (15 for /t'k'/ and 10 for /t'p'/) occur as derived oblique forms of nouns with clusters /t'k/ and /t'p/ (e.g. *bat'k'e* 'dad', dat.sg. from *bat'ka* 'dad' nom. sg.).[16]

Clusters with /t'/, on the other hand, are common across word boundaries, as shown in (2.59).

(2.59)	word boundary	
	/t/	/t'/
	ma[t#nʲ]ikolaja	ma[tʲ#nʲ]ikolaja
	Nikolaj's foul language	*Nikolaj's mother*
	ma[t#sʲ]ergeja	ma[tʲ#sʲ]ergeja
	Sergej's foul language	*Sergej's mother*
	ma[t#kʲ]irilla	ma[tʲ#kʲ]irilla
	Kirill's foul language	*Kirill's mother*

Thus, the contrast between /t/ and /t'/ before palatalized consonants, although formally possible, is almost non-existent lexically.

2.2.2.2.2.4 Other factors

The pattern described above reflects the Russian pronunciation of relatively younger Russian speakers. This differs in some ways from the prescribed norm and the earlier variety (the 'Old Moscow pronunciation') (Avanesov 1972: 145-146). In all of the attested patterns of spoken Russian the palatalized consonants emerge as the more likely outcome of neutralization, especially if the following consonant is homorganic. There are also various degrees of cohesion between hetero-organic consonants (Timberlake 1993: 830). A coronal stop is more likely to agree in palatalization with a following labial than a velar, and a labial is more likely to agree with a velar rather than coronal. The statement in (2.60a) implies that the most likely environment for the outcome of the neutralized [t'] is the following homorganic /t'/, followed by the hetero-organic /p'/ and then /k'/. Note that the order of preference for the labial is not exactly the same (2.60b).

(2.60) Place of C_2

a. Coronal /t/ vs. /t'/, [t']: _k' > _p' > _t'

b. Labial /p/ vs. /p'/, [p']: _t' > _k' > _p'

The pattern of distribution in homorganic clusters is found mainly if the following segment is a stop (/t'/) or a nasal (/n'/). The restriction against the sequence plain-palatalized (*CC') is less commonly found when /t/ is followed by the fricative /s'/ or the liquid /l'/. It never holds when C_2 is /r'/. The implicational statement based on the relative of application of the restriction is shown in (2.61).

(2.61) Manner of C_2

Coronal /t/ vs. /t'/, [t']: _r' > _l' > _s' > _j > _n' > _t'

In addition, the frequency of application of the described restrictions depends on a number of morphological, prosodic and lexical factors. Some of these factors are summarized below (based on Avanesov 1972: 145-165).

First, a consonant tends to agree in palatalization with the following segment more often if the cluster is within a stem, especially medially, than across a prefix + stem boundary. Agreement is less common if the cluster occurs across the preposition + word boundary (To the best of my knowledge, there is no record of segments agreeing across word boundaries).

Second, the constraint *CC' holds more often in stressed syllables rather than in unstressed ones. Third, the restriction is less common in infrequent, 'learned' lexical items and unassimilated loans than in common words. And finally fourth, the restriction applies more often

---

[16] There are also 3 words with the cluster /t'm'/ (*-t'm'-* and *-d'm'-*) derived from /t'm/.

in forms that exhibit no morphological alternations (i.e. where the palatalized C_2 does not alternate with the plain one).

These factors will be relevant for our later discussion. It is important that although relatively subtle in perceptual terms, these phonotactic differences cannot simply be attributed to universal phonetic implementation, since they are employed in different ways in a number of varieties of Standard Russian and dialects.

2.2.3 Conclusion

The examination of the distribution of Russian stops in several environments reveals severe restrictions on the occurrence of the plain-palatalized contrast. Some of the restrictions are categorical in nature, while others are tendencies involving phoneme frequencies, and morphological and lexical factors. The contrast is most robust in onset position, and highly restricted before consonants. In the latter context the palatalized (or plain) segments are often conditioned by the presence or absence of morpheme boundaries, alternations, and affiliation with some morphological and lexical classes.

An account of the facts of palatalization should be able to explain both categorical and gradient distributional effects.

Chapter 3. An acoustic study: Bursts and transitions

3.0 Introduction

In Chapter 2 I established the facts about the environments in which the plain-palatalized contrast is the best and the least well supported and about which consonants are most and least likely to undergo the neutralization process.

In this chapter I begin my investigation of the phonetic factors which underlie this distribution. The chapter focuses on the acoustics of this contrast in Russian in coda positions.

It is reasonable to ask the following questions: How does a language maintain a marked phonemic contrast in a marked environment? Is there something in the acoustics of Russian palatalized consonants in coda that makes them resistant to neutralization and are Russian listeners sensitive to this additional acoustic information? The examination of acoustics will provide us with an important point of reference for our later experiments in Chapters 4, 5, and 6.

I begin with a review of previous work on the acoustics of Russian plain and palatalized stops. Then I present results of a series of acoustic experiments focusing on the plain-palatalized contrast in coronals in coda environment. Recall that the distinction between /t/ and /t'/ in Russian is the least susceptible to neutralization.

3.1 Previous work on Russian plain and palatalized consonants

Previous acoustic studies or descriptions of Russian plain and palatalized consonants (Halle 1959, Kuznetsova 1965, Matusevich 1976, Bondarko 1981, Purcell 1979, Bolla 1981) determined that palatalized consonants are characterized by a high second formant (F2): lowering it during the CV transition and raising it during the VC transition. Most of these works focused on initial and medial prevocalic consonants, rather than final and medial preconsonantal segments. Thus, there is relatively little data on VC transitions in the latter contexts. While Halle 1959 did not find consistent differences between the spectral characteristics of Russian plain and palatalized bursts (150), Bolla 1981 described the palatalized coronal stop as having dominance at higher frequencies, as opposed to the relatively flat spectrum of the plain coronal stop (119-120). Some durational characteristics are mentioned in Kuznetsova 1969 and Bolla 1981, sometimes without reference to the specific environment.

3.2 Acoustic experiments:

/t/ vs. /t'/ in final and preconsonantal position

What acoustic parameters distinguish plain and palatalized coronal stops in coda environments? In this section I look at the differences between plain /t/ vs. palatalized /t'/ in final and preconsonantal coda positions. Experiment 1 deals with durational characteristics of plain and palatalized stops (preceding vowel, closure and burst duration). Experiment 2 focuses on the F2 formant patterns distinguishing the contrast.

3.3 Materials and procedure

Six native speakers of Russian, three males (Speakers 1, 4, 6) and three females (Speakers 2, 3, 5), participated in the experiment.[17] Test words presented to the speakers consisted of (near-)minimal word pairs containing phonemes /t/ and /t'/ in word final position (V_#, where V = /a/) and in medial preconsonantal environment (V_C; where V = /a/, C = /k/) (Table 3.1). In near-minimal pairs stress and the plain-palatalized quality of the preceding consonant was controlled for.

[17] The subjects are all speakers of Standard Russian in its several territorial variants: Speaker 1 (Lipetsk/Moscow), Speaker 2 (Moscow), Speaker 3 (Moscow), Speaker 4 (Arkhangelsk/Moscow), Speaker 5 (Vilnius, Lithuania), and Speaker 6 (Perm'). Speakers 1 and 4 have lived for more than 15 years in Moscow. Speakers 5 and 6 (the author) are graduate students in the University of Toronto Linguistics Department.

Table 3.1. (Near)-minimal pairs contrasting /t/ and /t'/ used in Experiment 1

	/t/		/t'/	
1.	ma[t]	*foul language*	ma[t']	*mother*
2.	mja[t]	*poached*	mja[t']	*to poach*
3.	snja[t]	*taken off*	snja[t']	*take smth. off*
4.	pa[t]	*stalemate*	spa[t']	*to sleep*
5.	pja[t]	*heel,* gen.pl.	pja[t']	*five*
6.	zhena[t]	*married*	gna[t']	*to drive*
7.	ka[t]ka	*pail*	Ka[t']ka	*Katja,* familiar
8.	lopa[t]ka	*scapula*	ba[t']ka	*dad,* familiar
9.	ma[t]ka	*womb*	Na[t']ka	*Nadja,* familiar
10.	pja[t]ka	*heel*	sja[t']ka	*sit down*
11.	vzja[t]ka	*bribe*	dja[t']ka	*uncle,* familiar
12.	smja[t]ka	*half-done eggs*	tja[t']ka	*dad,* familiar

Three repetitions of the stimuli were randomized and embedded in a carrier phrase [Skazhite ___] ("Say __"). Text words were utterance-final in order to ensure that the final bursts of stops were not affected by a following segment.[18] The sentences, interspersed with unrelated filler words, were presented in lists in Russian orthography, which exhibits the contrast. Speakers were recorded in a quiet room using a Marantz tape-recorder. Before recording, speakers practised reading a few randomly chosen test sentences to familiarize themselves with the materials. Materials were read at a comfortable speed throughout the recording session.

3.4 Analysis

For each speaker, only the second production was chosen for analysis. However, in a few cases, when a signal was not clear enough, the third or the first productions were used. A total of 72 tokens (12 words × 6 speakers) were digitized at a sampling rate of 11 kHz with 16-bit resolution, and stored as files to be processed by Signalyze 3.2 (http://www.agoralang.com/signalyze.html).

For a target word, the following parameters were measured: duration (ms) of the preceding vowel, consonant closure, consonant release burst; and frequency (Hz) of the second formant (F2) of the preceding vowel. Release burst spectra were also analyzed using MacQuirer (Scicon

Corp., http://www.sciconrd.com/). Duration measurements were done on the basis of both wide-band spectrograms and waveforms, with additional reference to zero-crossing and envelope. Vowel onset was considered to be the onset of the first formant, which corresponds to the onset of periodicity of the waveform. Vowel offset was taken as the offset of the second formant (on the spectrogram). Closure duration was defined as an interval between the vowel offset and release burst. Consonant-to-vowel and burst-to-closure ratios were calculated in order to control for different speech rates. Measurements of F2 of the vowel and burst parameters were done using LPC (Linear Predictive Coding) spectra, with reference to wide-band spectrograms and spectra, and zero crossing. For all tokens F2 was measured at the onset of the consonant (F2 final). Additional measurements were made at other points in time for selected words. The choice of the above mentioned measurements was suggested by findings from previous studies of palatalized consonants (See section 3.1) and stops in general.

The collected data, averaged across speakers, were analyzed in separate analyses of variance (ANOVA). The design of each analysis will be described for each experiment separately.

3.5 Experiment 1: Duration

3.5.1 Setup

To examine the duration parameters I ran six separate ANOVAs. Each test involved two between-item factors, Position (final and preconsonantal) and Following Consonant (plain and palatalized) and one dependent variable (consonant duration, vowel duration, consonant-to-vowel ratio, closure duration, burst duration, burst-to-closure ratio). In the case of vowel duration, the effect of the Preceding Consonant (plain and palatalized) was also analyzed separately as a between-item factor. Separate ANOVAs were run to determine inter-speaker

[18] The test words in the context before vowel-initial words (e.g. *Skazhite mat' op'at'*) were read by Speakers 4 and 5. Due to the time limits they were not analyzed.

variability in terms of all of these parameters. They involved a between-item factor, Speaker (six levels) and one dependent variable (consonant duration, vowel duration, consonant-to-vowel ratio, closure duration, burst duration, and burst-to-closure ratio). In addition to the ANOVAs, post hoc tests were run in order to determine the source of the interaction between factors or levels of factors.

3.5.2 Results

Table 3.2 presents means and standard deviations for several duration parameters. The statistical significance of each parameter is discussed below.

Table 3.2. Means (standard deviations) for acoustic parameters of /t/ and /t'/

Parameter	***Final***		***Preconsonantal***	
	/t/	*/t'/*	*/t/*	*/t'/*
Consonant duration, ms	181 (18)	186 (16)	84 (10)	94 (14)
Vowel duration, ms	149 (23)	150 (23)	113 (14)	120 (9)
Consonant-to-Vowel ratio	1.23 (.21)	1.26 (.21)	.76 (.10)	.78 (.10)
Closure duration, ms	100 (17)	71 (15)	59 (8)	43 (8)
Burst duration, ms	72 (29)	115 (18)	26 (9)	45 (20)
Burst-to-Closure ratio	.84 (.20)	1.71 (.60)	.45 (.18)	1.23 (.37)

3.5.2.1 Consonant duration

As shown in Table 3.3, plain and palatalized stops both finally and before consonants were not significantly different with respect to duration [$F(1,20) = 1.44$, $p = .24$]. At the same time the difference between final and preconsonantal positions was highly significant [$F(1,20) = 238.86$, $p < .001$]. There was no significant interaction.

Table 3.3. Consonant duration. Significant results are given in bold

Source of Variance	*DF*	*F*	*p*
Consonant (plain-palatalized)	1, 20	1.44	.243
Position (final-preconsonantal)	1, 20	**238.86**	**.000**
Consonant * Position	1, 20	.11	.747

The mean overall duration of final /t/ was 181 (18) ms and the duration of final /t'/ was 186 (16) ms. The mean values for preconsonantal /t/ and /t'/ were 84 (10) ms and 94 (14) ms respectively.

3.5.2.2 Vowel duration

The duration of the preceding vowel (Table 3.4) did not show any significant difference depending on the plain-palatalized following consonant [$F(1,20) = .33$, $p = .57$]. As in the previous case, Position was a significant factor: vowels before both consonants were shorter preconsonantally [$F(1,20) = 19.19$, $p < .001$]. No significant interaction was found.

Table 3.4. Vowel duration

Source of Variance	*DF*	*F*	*p*
Consonant (plain-palatalized)	1, 20	.33	.570
Position (final-preconsonantal)	1, 20	**19.19**	**.000**
Consonant * Position	1, 20	.18	.677

The mean vowel duration before final /t/ was 149 (23) ms and the vowel duration before final /t'/ was 150 (23) ms. The mean values before preconsonantal /t/ and /t'/ were 113 (14) ms and 120 (9) ms respectively.

The influence of the plain-palatalized quality of the *preceding* consonant (e.g. *m'at/m'at'* vs. *mat/mat'*; *p'atka/s'ad'ka* vs. *matka/Kat'ka*), was significant. Vowels were found to be longer if the preceding consonant was palatalized [$F(1,20) = 22.144$, $p < .0001$]. Thus, in final position the mean vowel duration after a plain consonant was 142 (22) ms before /t/ and 140 ms before /t'/. The vowel duration after a palatalized segment was 159 (24) ms before /t/ and 158 (23) before /t'/. In preconsonantal position the mean values after preconsonantal /t/ and /t'/ were 115 (13) ms and 104 (15) ms after a plain consonant and 125 (8) and 125 (10) after a palatalized one. No significant interaction of factors was found.

3.5.2.3 Consonant-to-vowel ratio

As shown in Table 3.5, the results for consonant-to-vowel ratio are in line with the previous findings: Consonant (/t/ vs. /t'/) was not significant [F(1,20) = .15, p= .70], but the main effect of Position was [F(1,20) = 50.40, p < .001].

The mean consonant-to-vowel ratio for final /t/ was 1.23 (.21) and the ratio for final /t'/ was 1.26 (.21). The mean values for preconsonantal /t/ and /t'/ were .76 (.10) and .78 (.10) respectively.

Table 3.5. Consonant-to-Vowel duration ratio

Source of Variance	*DF*	*F*	*p*
Consonant (plain-palatalized)	1, 20	.15	.704
Position (final-preconsonantal)	1, 20	**50.40**	**.000**
Consonant * Position	1, 20	.00	.971

3.5.2.4 Closure duration

Table 3.6. Closure duration

Source of Variance	*DF*	*F*	*p*
Consonant (plain-palatalized)	1, 20	**18.81**	**.000**
Position (final-preconsonantal)	1, 20	**44.88**	**.000**
Consonant * Position	1, 20	1.36	.258

Closure duration measurements revealed two significant effects (Table 3.6). First, the closure of palatalized /t'/ was shorter than that of plain /t/ [F(1,20) = 18.80, p < .001]. Position was also highly significant: F(1,20) = 44.88, p < .001. The interaction was not significant [F(1,20) = .001, p = .971].

The mean closure duration of final /t/ was 100 (17) ms and the closure duration of final /t'/ was 71 (15) ms. The mean values of preconsonantal /t/ and /t'/ were 59 (8) ms and 43 (8) ms respectively.

3.5.2.5 Burst duration

All final and preconsonantal stops were released. The burst duration for final and preconsonantal /t'/ was found to be significantly longer than that of /t/ in both environments [$F(1,20) = 14.32$, $p = .001$] (Table 3.7). Burst was also affected by Position. It was shorter before consonants than finally.

Table 3.7. Burst duration

Source of Variance	*DF*	*F*	*p*
Consonant (plain-palatalized)	1, 20	**14.32**	**.001**
Position (final-preconsonantal)	1, 20	**49.58**	**.000**
Consonant * Position	1, 20	1.95	.178

The mean burst duration of final /t/ was 72 (29) ms and the burst duration of final /t'/ was 115 (18) ms. The mean values of preconsonantal /t/ and /t'/ were 26 (9) ms and 45 (20) ms respectively.

3.5.2.6 Burst-to-closure ratio

The importance of closure and burst in contrasting plain and palatalized stops was confirmed by the burst-to-closure ratio data (Table 3.8). Both factors, Consonant [$F(1,20) = 28.77$, $p < .000$] and Position [$F(1,20) = 7.87$, $p = .011$], were significant.

Table 3.8. Burst-to-Closure ratio

Source of Variance	*DF*	*F*	*p*
Consonant (plain-palatalized)	1, 20	**28.77**	**.000**
Position (final-preconsonantal)	1, 20	**7.87**	**.011**
Consonant * Position	1, 20	.07	.798

The mean burst-to-closure ratio for final /t/ was .84 (.20) and the ratio of final /t'/ was 1.71 (.60). The mean values of preconsonantal /t/ and /t'/ were .45 (.18) and 1.23 (.37) ms respectively.

3.5.2.7 Inter-speaker variation

Although the individual differences between speakers were not the focus of the experiment, they were still considered. The analyses of variance showed that the effect of Speaker was not significant for all durational parameters (Consonant [$F(5, 18) = .166$; $p = .972$], Vowel [$F(5, 12) = .681$; $p = .644$], Consonant-to-Vowel Ratio [$F(5, 12) = .654$; $p = .662$], Closure [$F(5, 12) = .112$; $p = .988$], Burst [$F(5, 12) = .388$; $p = .850$], Burst-to-Closure Ratio [$F(5, 12) = .231$; $p = .944$]).

3.5.3 Discussion

The findings showed that plain and palatalized coronal stops differed primarily in closure and burst duration. This held true for these consonants in both final (Figure 3.1) and preconsonantal positions (Figure 3.2). While the plain closure was usually longer than the palatalized one (difference of 30 ms), the palatalized burst was longer than the plain one (difference of 35 ms). Both palatalized and plain coronal stops were strongly affricated: their burst was substantially longer than closure. This was especially true for the palatalized stop /t'/, whose frication period was on average 1.6 times longer than its closure.

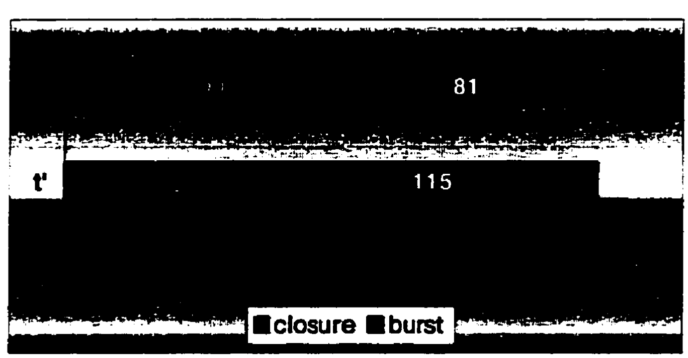


Figure 3.1. Mean closure and burst values for /t/ and /t'/ in final position (in ms).

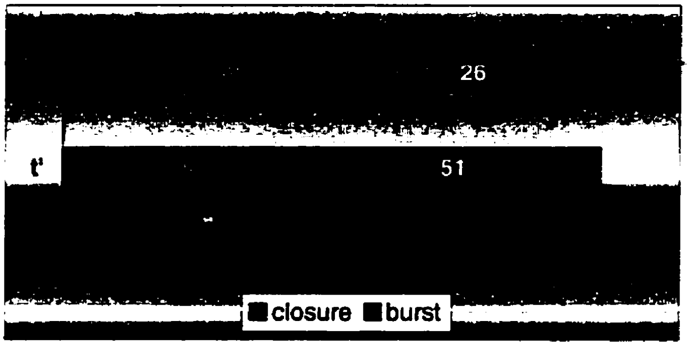


Figure 3.2. Mean closure and burst values for /t/ and /t'/ in preconsonantal position (in ms).

Preconsonantal palatalized /t'/ differed significantly from final /t'/ in the duration of the preceding vowel (difference of 30 ms), as well as in closure and burst duration (differences of 28 and 64 ms respectively), (Figure 3.3). The same was true for the plain /t/ in the two positions (Figure3.4). The difference between final and preconsonantal environments was 36 ms for the vowel, 41 ms for the closure, and 55 ms for the burst. Of the three components, burst was the most affected by position: the burst of /t'/ before a consonant was 44% of the final burst and the burst of /t/ was only 32% of the burst in the final position.

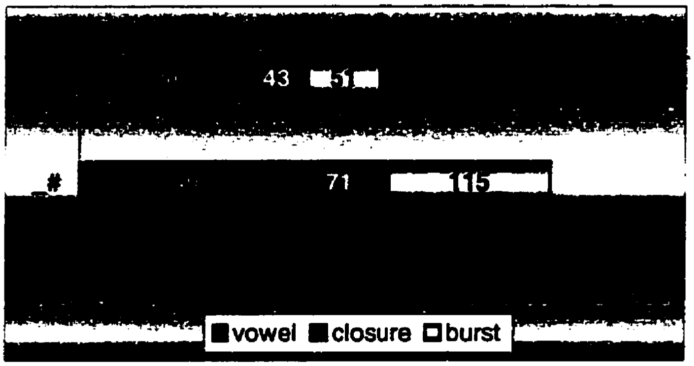


Figure 3.3. Mean duration of vowel, closure, and burst for /t'/ in final and preconsonantal positions (in ms).

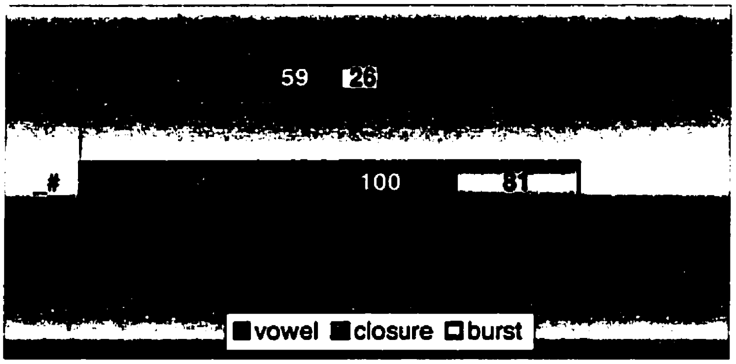


Figure 3.4. Mean duration of vowel, closure, and burst for /t/ in final and preconsonantal positions (in ms).

Both segments were less affricated (i.e. characterized by a smaller burst-to-closure ratio) in preconsonantal position compared to the final environment. Interestingly, however, the contrast between /t/ and /t'/ in terms of their burst-to-closure ratio remained almost invariant (Figure 3.5).

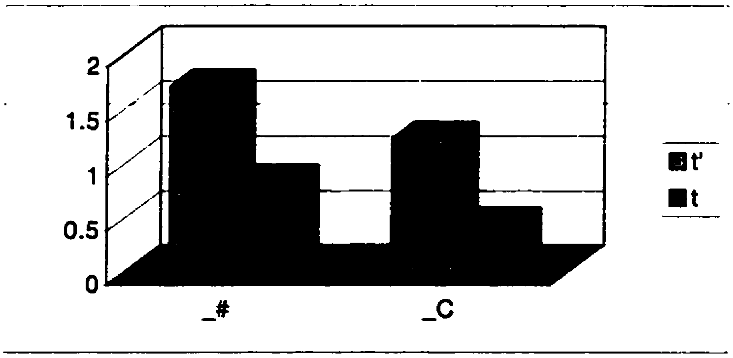


Figure 3.5. Burst-to-closure ratio in final and preconsonantal positions.

The presence of release in Russian, regardless of position, was noted in Jones & Ward 1969 and Bondarko 1977. Note that a stop burst is usually described in the literature as having a short duration (5 to 40 ms) and the burst-to-closure ratio from .1 to .4 (Kent & Read 1992: 110). Moreover, in many languages, e.g. English, the burst release in coda is optional.

Our findings of palatalized burst and closure duration in coda also support the results based on one speaker reported in Kuznetsova 1969 that show similar relations between preconsonantal (closure of 40 ms and burst of 76 ms) and final positions (closure of 61 ms and burst of 116 ms) (113). However, the usual assumption that palatalized stops have longer overall duration is not supported by our data (Bolla 1981, Bondarko 1981). It is likely that this holds for onset consonants only.

3.6 Experiment 2: Second formant of the vowel

3.6.1 Setup

Experiment 2 was designed to determine whether the plain-palatalized contrast was signaled by the quality of the preceding vowel, particularly by the VC transition. Recall that vowel duration did not differ significantly between the two contexts (plain or palatalized following consonant) (See Table 3.4).

I limited the analysis of VC transition to the second formant (F2) frequency. I examined the second formant values both at the offset of the vowel, F2 ending and at four points during the vowel. The first analysis was based on averages of 10 words (pairs 1, 2, 3, 4, 5, 7, and 8 from Table 3.1) per speaker, while the second one was limited to the minimally contrastive words with initial labials (final position) and velars (preconsonantal position).[19]

To examine F2 ending I performed an analysis of variance with two between-item factors, Position (final and preconsonantal) and Consonant (plain and palatalized) and a dependent variable (F2 value). In the analysis of F2 throughout the vowel I ran five separate ANOVAs for each of the two positions. The analysis of the final position included two between-item factors,

[19] The inclusion of all words would make the comparison more complicated due to additional factors of preceding consonants of different places of articulation.

Following Consonant and Preceding Consonant (both plain and palatalized) and a dependent variable (F2 values at five points). The analysis of preconsonantal position was limited to the environment after a plain consonant, and thus had one between-item factor, Following Consonant (plain and palatalized) and a dependent variable (F2 values at five points).

3.6.2 Results

3.6.2.1 F2 ending

Means and standard deviations of the second formant values at the point before the closure of the final consonant are given in Table 3.9. The average difference between plain and palatalized F2 was from 459 Hz before consonants to 557 Hz finally.

Table 3.9. Mean and standard deviation (in parenthesis) values for F2 before /t/ and /t'/

Parameter	***Final***		***Preconsonantal***	
	Plain	*Palatalized*	*Plain*	*Palatalized*
F2 final, Hz	1452 (216)	2009 (178)	1467 (227)	1926 (169)

The ANOVA results revealed that the difference between plain and palatalized stops with respect to F2 was significant [$F(1,20) = 39.16$, $p < .001$] (Table 3.10). Position was of no importance here [$F(1,20) = .33$, $p = .56$]: the formant properties distinguishing the consonants remained fairly stable.

Table 3.10. F2 ending

Source of Variance	*DF*	*F*	*p*
Following Consonant (plain-palatalized)	1, 20	**39.16**	**.000**
Position (final-preconsonantal)	1, 20	.33	.570
Consonant * Position	1, 20	.36	.558

3.6.2.2 F2 throughout the vowel

To compare the F2 values throughout the vowel, three minimal pairs were chosen: *mat* vs. *mat'*, *m'at* vs. *m'at'*, and *kadka vs. Kat'ka*. The first four words differ minimally from each other, in terms of the preceding (labial) and/or following (coronal) plain or palatalized consonant. Since there are no lexical items that would show the same minimal contrast in the preconsonantal position, only two words contrasting in the following coronal consonant were used. Measurements were made at five points in time: at the onset of the vowel, at the onset of the consonant closure, and at three additional points within the vowel: 30 ms, 60 ms, 90 ms before the closure. These points divided a given vowel into almost five equal intervals (the mean vowel duration was about 120 ms), and allowed us to compare the dynamics of F2 in before plain and palatalized stops. Since the analyses of F2 in final and preconsonantal positions differed in the number of factors, they are presented separately.

3.6.2.3 Final position

Table 3.11 presents means and standard deviations of F2 values at 5 points in time depending on the plain-palatalized quality of the preceding and following consonant.

Table 3.11. Mean and standard deviation (in parenthesis) values for F2 (Hz), final position

	F2 starting	F2 ending minus 90 ms	F2 ending minus 60 ms	F2 ending minus 30 ms	F2 ending
CaC	1028 (89)	1230 (171)	1291 (137)	1362 (221)	1485 (218)
C'aC	1949 (73)	1706 (188)	1626 (175)	1504 (197)	1473 (225)
CaC'	1089 (109)	1335 (88)	1559 (175)	1804 (179)	1962 (179)
C'aC'	2045 (169)	1877 (196)	1857 (219)	1943 (172)	2019 (182)

As we see in Table 3.12, the Preceding Consonant was a significant factor at the first three points: [$F(1,20) = 394.141$, $p < .001$], [$F(1,20) = 55.960$, $p < .001$], and [$F(1,20) = 19.398$, $p = .003$]. The Following Consonant significantly affected F2 values at the last three points: [$F(1,20) = 12.197$, $p = .002$], [$F(1,20) = 31.177$, $p < .001$], and [$F(1,20) = 38.459$, $p < .001$]. No interaction between the factors was found.

Table 3.12. F2 at five points in time, final position (DF 1,20 for all)

	F2 starting	F2 ending -90 ms	F2 ending -60 ms	F2 ending -30 ms	F2 ending (0 ms)
Following C					
F	2.75	4.12	**12.20**	**31.18**	**38.46**
p	.113	.056	**.002**	**<.001**	**<.001**
Preceding C					
F	**394.14**	**55.96**	**19.40**	3.16	.75
p	**<.001**	**<.001**	**<.001**	.091	.787

Figs. 3.6 and 3.7 plot F2's of the vowel /a/ in the words *mat* vs. *mat'* and *m'at* vs. *m'at'*, averaged across six speakers. We can see from the figures that vowel F2 (CV transition) was affected by the preceding consonants at the beginning of its trajectory. F2 in Figure 3.6 starts at around 1000 Hz after a plain labial. Then it stays relatively level from point 2 to point 4, after which it rises further to 1500 Hz if the following consonant is plain. If the following consonant is palatalized, F2 gradually moves up during almost the entire vowel, ending at around 2000 Hz. The contrast between plain and palatalized F2's is observed during the most of the vowel, and is particularly obvious in the second part of the vowel (roughly, the last 60 ms).

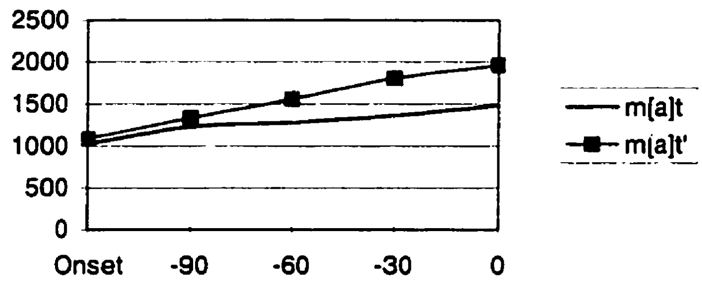


Figure 3.6. Mean F2 values for [a] in the words ma[t] and ma[t']. Measurements were taken at the onset and offset of the vowel, 30 ms, 60 ms, and 90 ms prior to the offset.

Figure 3.7 shows F2 of the vowel [a] *after* a palatalized /m'/. While in the context before a plain consonant F2 gradually goes down to around 1500 Hz, in the environment of the following /t'/ the formant starts at around 2000 Hz and stays fairly high (around 1800-2000 Hz). The resulting sound has been noted to be similar to the low front vowel [æ], rather than to the prototypical [a] (Matusevich 1976: 72-74).

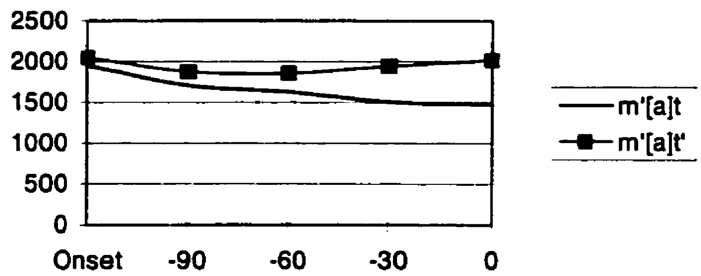


Figure 3.7. Mean F2 values for [a] in the words m'at and m'at'.

For both environments the major contrast between the following plain and palatalized consonant was in the second half of the vowel. The plain coronal stop /t/ was characterized by F2 in mid-frequencies and the palatalized /t'/ showed F2 values in the high frequencies. We can also say that a transition to a palatalized consonant was longer and steeper than that to a plain consonant.

3.6.2.4 Preconsonantal position

Table 3.13 presents means and Standard deviations of F2 values at 5 points in time in preconsonantal position depending on the plain-palatalized quality of the following consonant.

Table 3.13. Mean and standard deviation (in parenthesis) values for F2 (Hz), preconsonantal position

	F2 starting	F2 ending minus 90 ms	F2 ending minus 60 ms	F2 ending minus 30 ms	F2 ending
CaC	1518 (167)	1447 (170)	1440 (183)	1492 (183)	1468 (208)
CaC'	1598 (102)	1597 (81)	1686 (58)	1822 (129)	1943 (152)

As in the previous case, we found a significant influence of the following consonant on the vowel, particularly on its second half (Table 3.14). We can observe the significant difference at the last three points: [$F(1,20) = 9.892$, $p = .010$], [$F(1,20) = 13.005$, $p = .005$], and [$F(1,20) = 20.444$, $p = .001$]. No interaction between the factors was found.

Table 3.14. F2 at five points in time, final position (DF 1,20 for all)

	F2 starting	F2 ending minus 90 ms	F2 ending minus 60 ms	F2 ending minus 30 ms	F2 ending
Following C					
F	1.02	3.80	**9.89**	**13.00**	**20.44**
p	.336	.080	**.010**	**.005**	**.001**

Figure 3.8 illustrates the trajectories of F2 before plain and palatalized coronal stops. The preceding consonant is a plain velar. This explains a higher starting F2 (around 1500 Hz), compared to the preceding plain labial in Figs. 3.6 and 3.7. Other than that F2 in the preconsonantal environment exhibits a pattern similar to the F2 trajectory in the final position.

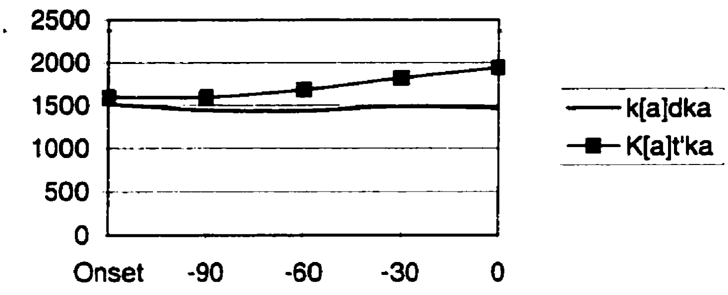


Figure 3.8. Mean F2 values for [a] in the words ka[t]ka and Ka[t']ka.

3.6.2.5 Inter-speaker variation

To determine the variability of F2 across speakers I ran a separate ANOVAs with a between-item factor, Speaker (six levels) and one dependent variable (F2 Ending). While the data show some variation across speakers between the plain and palatalized F2 trajectories, the differences were determined to be not significant [$F(1,5) = 1{,}571$, $p = .2185$].

3.6.3 Discussion

The results show that the plain and palatalized distinction in coda is not signaled solely by the consonants themselves, but also by the preceding vowels, particularly by the second half of the vowel. The difference in the ending value of F2 was almost constant for both final and preconsonantal environments (about 300 Hz). F2 trajectories varied depending on the preceding consonant. Given the results, the first parameter, F2 ending, is likely to be the main correlate of the contrast.

The findings conform to previous studies that note the importance of F2 transition in encoding the contrast (e.g. Kuznetsova 1965, Bondarko 1981, Purcell 1979). At the same time our data suggest a higher value for F2 before palatalized coda stops (2000 Hz) than was reported in Matusevich 1976: 72-74 and Bondarko 1977: 146 (1500 Hz). Note also that we did not find that F2 of the palatalized CV transition was higher than that of the palatalized VC transition (Table 3.11), as mentioned in these sources. In this respect our results are closer to the findings for VC transitions in VCV sequences in Purcell 1979.

Note that the mean F2 value before the Russian plain /t/ (1450 Hz) is much lower than the mean F2 of an English /t/ (1800 Hz (Kent & Read 1992: 136-138; Johnson 1997: 136)). This fact may be explained by the dental articulation of the Russian coronal or possibly by its secondary velar constriction. Additional acoustic and articulatory analysis is necessary to answer the question of just what accounts for the difference.

3.7 Experiment 3: Burst spectra

3.7.1 Setup

This experiment dealt with the spectral characteristics of burst of fricative noise in distinguishing the plain-palatalized contrast. I examined the spectrograms and FTT/LPC spectra of burst of minimal pairs *mat* vs. *mat'* and *kadka vs. Kat'ka* pronounced by each speaker. The spectra were taken at the onset of the fricative noise and in the middle of the noise period (for final position only).

3.7.2 Results

The analysis of spectrograms of final /t/ and /t'/ (Figure 3.9) revealed that the plain burst could be characterized as having an abrupt onset of noise (or short rise time) at a wide range of frequencies from 300-600 Hz to 5000 Hz and higher. The areas of high energy were found around 1700-1800 Hz and above 3300 Hz (for some speakers) or 4200 Hz (for others). On the other hand, the spectrograms of palatalized burst showed a slow, fricative-like, onset of noise (or longer rise time) at high frequencies, above 3700-4000 Hz, followed by high energy areas at around 2600-2900 Hz, and particularly at around 3800-4500 Hz. Later the latter peak lowered to 3500 Hz. For some speakers there was a minor peak at around 1800 Hz.

Omat-mat'-Spectrogram.PICT

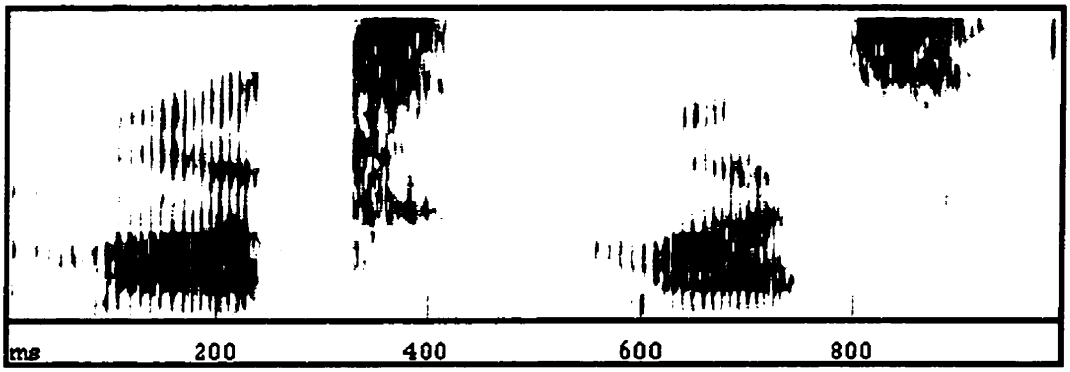
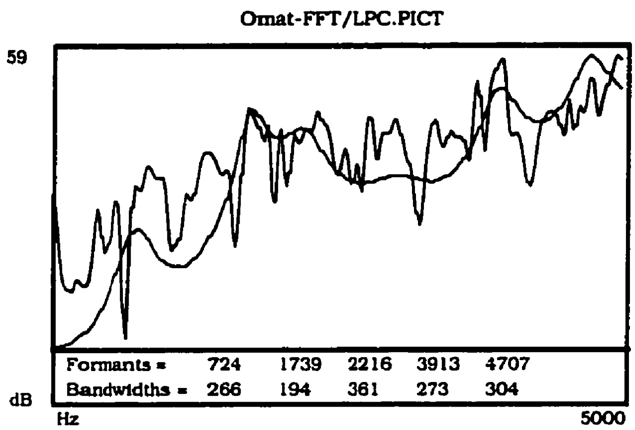


Figure 3.9. A spectrogram of the minimal pair words *ma[t]* and *ma[t']*
as pronounced by Speaker 1.

Our analysis of FFT/LPC spectra (Figure 3.10) shows that the burst of plain /t/ has peaks at frequencies from 700-800 Hz to 4500-4800 Hz with a major mid-frequency peak at around 1500-1900 Hz. For some speakers (Speakers 1 and 2) this peak is accompanied by a prominent high frequency peak in the 4200-4700 Hz range. In general the plain /t/ spectrum can be characterized as diffuse and flat (with no or little spectral tilt). The palatalized spectra showed a strong high-frequency dominance at 4500-4800 Hz and higher with much lower amplitude at lower frequencies (the only exception was Speaker 4 whose spectrum had a lower main peak at 2700 Hz). The palatalized burst can be characterized as less diffuse with a strong rising tilt.

Both the spectra of /t/ and /t'/ in the second half of the frication period showed the patterns similar to the ones described above, though at somewhat lower frequencies. It seems that during this period the frication noise was accompanied by strong aspiration.



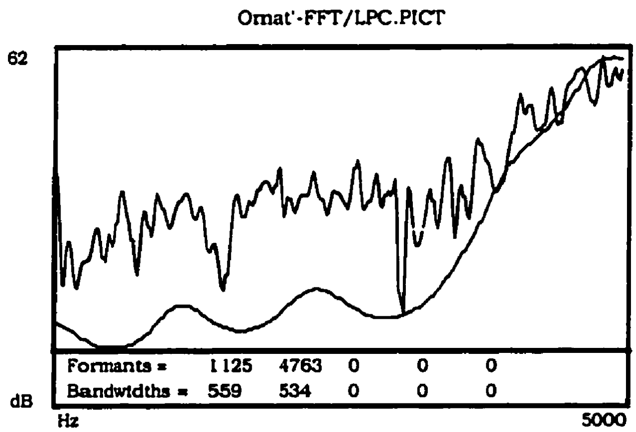


Figure 3.10. FFT/LPC spectra of bursts of plain /t/ and palatalized /t'/ as pronounced by Speaker 1.

While the duration of the burst period in preconsonantal environment is much shorter (Figure 3.11) for both plain and palatalized coronals, their spectral characteristics were found to be similar to the ones in final position.

3.7.3 Discussion

The results of the experiment showed that plain and palatalized stops differ in spectral characteristics of their frication periods. The diffuse and relatively flat spectrum of /t/ is in contrast with the less diffuse rising spectrum of /t'/. These findings are similar to the results reported in Bolla 1981: a wide range noise of /t/ vs. the most intensive noise around 4200-6800

Hz for /t'/ (120-121). The presence of aspiration in the frication period of final stops was noted by Bondarko 1977:95, 1981: 132.

Okadka_kat'ka-Spectrogram.PICT

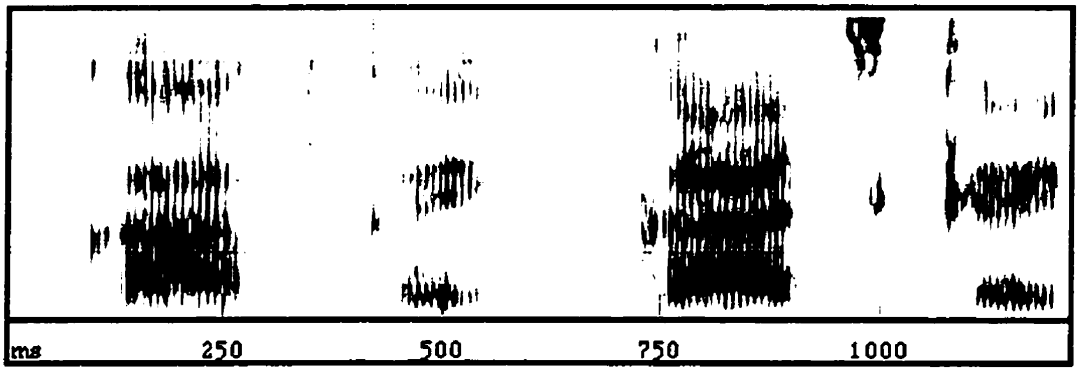


Figure 3.11. A spectrogram of the minimal pair words ka[t]ka and Ka[t']ka (pasted together) as pronounced by Speaker 1.

3.8 Summary

The contrast between the plain /t/ and palatalized /t'/ in final position is associated with several acoustic parameters, both spectral and temporal (Table 3.15): the duration of closure and burst, the quality of the preceding vowel, particularly the value of F2 in the second half of the vowel and its slope, and the quality of burst. Note that the VC transition parameters are limited to the context after /a/ preceded by a plain consonant and the preconsonantal burst is characteristic of the environment before /k/.

Table 3.15. Acoustic parameters associated with the contrast /t/-/t'/ (final position)

Segment (Coda position)	*VC transition*		*Closure*	*Burst*	
	F2 /a/ (Hz)	F2 ending (Hz)	Silence duration (ms)	Noise	Noise duration (ms)
/t'/	1300	2000	71/43*	diffuse flat	115/51*
/t/	1300	1450	115/59*	less diffuse rising	81/26*

* The second value is given for preconsonantal position.

Do all these parameters contribute to a perceptual correlate of the plain and palatalized coronal stops? What is the relative importance of these components? What acoustic information about the two segments is relevant for a listener and what acoustic events may be ignored?

We have found that the spectral differences of burst and vowel transitions are more consistent across the two positions than the temporal differences, closure and burst duration. Given this, we should expect that the former parameters would be more salient cues to the contrast than the latter ones. That is, removing a burst or reverting a burst or vowel transition should lead to shifting the phonemic boundary between the two segments. On the other hand, changing the duration of closure or burst should have a less significant effect on the identification of the segments. Moreover, reverting the quality of the first half of the vowel should have even lesser effect. These predictions will be tested in the following chapter.

Chapter 4. Acoustic characteristics and perception

4.0 Introduction

In this section we will examine the perceptual significance of the established acoustic parameters in identification of the plain-palatalized contrast between /t/ and /t'/ in final and preconsonantal positions. These are the spectral characteristics of burst and VC transition, as well as temporal factors, duration of burst and closure.

4.1 Previous work on Russian plain and palatalized consonants

Studies of perception of Russian plain and palatalized consonants are less numerous than the acoustic studies. Derkach 1975 examined the perception of synthetic plain and palatalized fricatives in the intervocalic environment. He found that the CV transition (both F1 and F2) was a major cue to the plain-palatalized contrast. Some studies referred to in Bondarko 1981 tested the perceptual cues of CV, VC transitions, and burst. The results showed that listeners could reliably identify initial plain and palatalized stops based on CV transitions. The identification of final stops was based on VC transition and was much less reliable. On the other hand, they determined that burst was an important cue to the final palatalized stop. Unfortunately, I have no access to the detailed descriptions and results of the experiments.

4.2 Procedure

Twelve native speakers of Russian were involved in this experiment, including most of the subjects used in the production experiments (Speakers 1, 2, 4, 5, and 6).[30] None of the listeners had any known hearing disorders. All subjects were tested individually.

Test words consisted of minimal pairs pronounced by all speakers (or by Speaker 1 only for experiment 3). They were taken from the phrases examined in acoustic experiments and excised from their original carrier phrases. The stimuli, organized into three blocks, were presented to listeners in random order with a two second response-to-stimulus interval, using the program PsyScope (Cohen, MacWhinney, Flatt, and Provost 1993). Listeners pressed one key if they heard a plain /t/ and the other key if they heard a palatalized /t'/. Response times (RT) were also measured. All of the stimuli were presented twice. The mean identification rates (proportion correct) were based on the averages of two tokens per word from twelve listeners. The results were further analyzed using repeated-measures ANOVAs, described separately for each experiment.

There were two experiments: Experiment 1 examined the role of the palatalized burst both finally and preconsonantally in the context of a preceding /a/. Experiment 2 explored the relative salience of various acoustic parameters as perceptual correlates of plain and palatalized coronal stops.

4.3 Experiment 1: Removing palatalized burst

The goal of this experiment is to determine whether removing a palatalized burst influences subjects' identification of the segment.

4.3.1 Materials and analysis

I considered two minimal pairs (Table 4.1) contrasting /t/ and /t'/ in two positions.

Table 4.1. Minimal pairs contrasting /t/ and /t'/

/t/		/t'/	
ma[t]	*foul language*	ma[t']	*mother*
ka[t]ka	*pail*	Ka[t']ka	*Katja,* familiar

There were three levels of the Burst factor: plain /t/, and two variants of words with palatalized /t'/: one with a complete burst, and the other with the burst removed. This yielded a total of 864 tokens (6 stimuli × 6 Speakers × 2 repetitions × 12 listeners).

[20] In all perceptual experiments the subjects-linguists (Subjects 5 and 6) showed results consistent with the average for the group (Rates .73 and .72 respectively; the mean identification rate is .72 (.02)).

In the analysis of response time (RT) only positive responses (i.e. all but errors in identification) were considered. RT of less than 50 ms and more than 2000 ms were excluded.[21]

The results, averaged across subjects and words, were evaluated by repeated-measures ANOVA with two within-subject factors: Burst (3 levels: plain, palatalized, and palatalized removed) and Position (2 levels: final and preconsonantal). The dependent variable corresponded to repeated measurements of identification at various factor combinations. The same design was used to evaluate response time. A separate repeated-measures ANOVA was run to determine the difference in terms of the preceding consonant (plain or palatalized) in the final position. The design included Preceding Consonant as a within-subject factor and dependent variables, repeated measurements.

4.3.2 Results

Means and standard deviations for identifications of stops are given in Table 4.2. The two positions are presented in rows and the three kinds of stimuli are given from left to right in columns. Identification scores are presented as a proportion of correct response, out of 1.

Table 4.2. Identification (out of 1) of the /t/-/t'/ contrast and response time (ms), final and preconsonantal positions. Mean and standard deviation values, averaged across subjects.

Identification	/t/	/t'/	/t'7
Final	.97 (.07)	1.00 (.00)	.92 (.08)
Preconsonantal	.91 (.08)	1.00 (.00)	.91 (.12)
RT			
Final	490 (157)	414 (145)	670 (148)
Preconsonantal	685 (173)	541 (192)	604 (183)

Table 4.3 presents means and standard deviations for response time in identification of the consonants in question. The statistical analysis suggests that removing the palatalized burst significantly affects the identification of /t'/ [$F(2, 22) = 10.75$; $p < .001$], while the difference in terms of position is not significant [$F(1, 11) = 2.24$; $p > .15$] (Table 4.3).

[21] The overall number of excluded tokens was 206 (in final position: 30 with /t/, 9 with /t'/, and 135 with /t'/ with

Table 4.3 Identification of the /t/-/t'/ contrast and response time

Identification	*DF*	*F*	*p*
Position (final-preconsonantal)	1, 11	2.24	.163
Burst (plain-palatalized-palatalized removed)	2, 22	**10.75**	**.001**
Burst * Position	2, 22	.89	.424
RT			
Position (final-preconsonantal)	1, 11	**8.39**	**.015**
Burst (plain-palatalized-palatalized removed)	2, 22	**22.61**	**.000**
Burst * Position	2, 22	**18.21**	**.000**

A Newman-Keuls test was performed to investigate the differences within the factor Burst. The difference between /t/ and /t'/ was significant ($p < .01$) as was the difference between /t'/ with complete vs. removed burst ($p < .001$). The difference between /t/ and /t'/ with no burst was not significant.

The analysis of measures of response time indicated significant differences in terms of Burst, Position, as well as a significant interaction between the two factors (See the interaction plot in Appendix, Figure 5.7). A Scheffé test showed that the significant RT differences were only between /t'/ and /t'/ with no burst ($p = .0064$). Plain /t/ did not differ either from /t'/, or from [t'ˀ].

In the final position, identifying a plain /t/ took longer than the palatalized /t'/; removing the burst led to longer processing of the segment. The pattern was somewhat different in the preconsonantal position. As in the previous case, it took longer to identify the plain stop and the palatalized unreleased stop than the released palatalized consonant.

removed burst; in preconsonantal position: 15 with /t/, 1 with /t'/, and 16 with /t'/ with removed burst).

Caution should be taken when considering the RT analysis. First, the stimuli with removed burst differ qualitatively from the unedited stimuli with plain and palatalized burst. It is likely that processing of a less naturally sounded signal would take more time.[22] Second, there was a difference between stimuli with final and preconsonantal stops. While in the final position /t/ and /t'/ were followed by a pause (*ma[t]#* and *ma[t']#*), in the preconsonantal position these consonants were followed by a sequence
-*ka*. (ka[t]ka and Ka[t']ka). This fact may have also influenced the difference in processing time between the two positions.

Whether the preceding consonant was plain or palatalized did not affect the identification of /a/ [$F(1,11) = 1.197$; $p = 0.297$]. No interaction was observed. Similar results were found for response time [$F(1,11) = .000$; $p = 0.998$].

4.3.3 Discussion

Figure 4.1 illustrates the findings. What we see here is that the palatalized /t'/ was always correctly identified, while the correct identification of the plain /t/ was less consistent, especially in the preconsonantal position. Removing the palatalized burst affected the identification of /t'/ in both positions, although the effect was very small. In these few cases /t'/ was identified as the plain /t/.

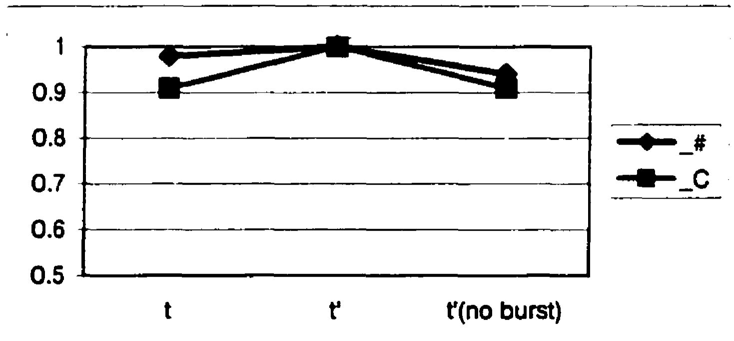


Figure 4.1. Identification of /t/, /t'/, and /t'/ with a removed burst in final and preconsonantal positions.

[22] Some listeners noted certain unnaturalness of some stimuli.

These results suggest that the palatalized burst plays some role in the identification of the segment. This was expected from our acoustic analysis. However, based on this experiment, the burst is not the most crucial cue to the segment, at least in the environment after /a/. The VC transition appears to be sufficient to support the identity of the sound most of the time.

The quality of the preceding consonant, and thus the slope of F2 (see section 3.4) did not play a significant role in the recognition of the consonants. This suggests that a high value of F2 rather than the slope is the cue to a palatalized consonant.

As expected, the temporal differences between the two positions (shorter duration of cues in the preconsonantal environment) do not seem to be a major factor. Even reduced in magnitude, the spectral cues to the contrast remain salient enough to be reliably recovered by a listener. The contrast may show sensitivity to the position in more adverse conditions (e.g. with added noise) or with increased rate of speech. This is to be considered in later chapters.

4.4 Experiment 2: Relative cue salience

In this experiment I investigate the relative importance of several acoustic components in cueing the contrast between plain and palatalized coronal stops in final and preconsonantal positions.

4.4.1 Materials and analysis

Two minimal pairs, *ma*[t] and *ma*[t'], *ka*[t]*ka* and *Ka*[t']*ka* (Figure 4.2) as pronounced by Speaker 1, were used to produce the stimuli.[23] Each word was segmented into four components: the first ('steady') part of the vowel with the preceding consonant, the second part of the vowel, closure, and burst (with the following part of the word, for ka[t]ka-Ka[t']ka only). Each of the components can be plain or palatalized. Recall that plain and palatalized bursts differ in their spectral and durational characteristics; the transitions (V2) differ primarily in their spectral

[23] Due to time constraints only one speaker was tested. This particular speaker was evaluated by two native Russian speakers as speaking 'Standard Russian' without a regional accent.

characteristics; the closures are distinguished solely by duration; the steady vowel segments differ only slightly at their end. The components were separated and spliced together in all possible VC combinations. The subjects had to determine whether the words included the sounds /t/ or /t'/.

The resulting 768 tokens (32 stimuli × 1 Speaker × 2 repetitions × 12 listeners) were collapsed across subjects and analyzed by a factors vs. repeated measures ANOVA with 5 within-subjects factors, Position, V-steady (V1) , VC transition (V2), Closure, and Burst. Each of the factors had two levels, plain and palatalized. A separate ANOVA with the same design was used in the analysis of response time. For the response time analysis both positive and negative responses were considered. Values less than 50 ms and more than 2000 ms were excluded (32 tokens).

4.4.2 Results

Identification scores and response time for final and preconsonantal positions are given in Tables 4.4 and 4.5.

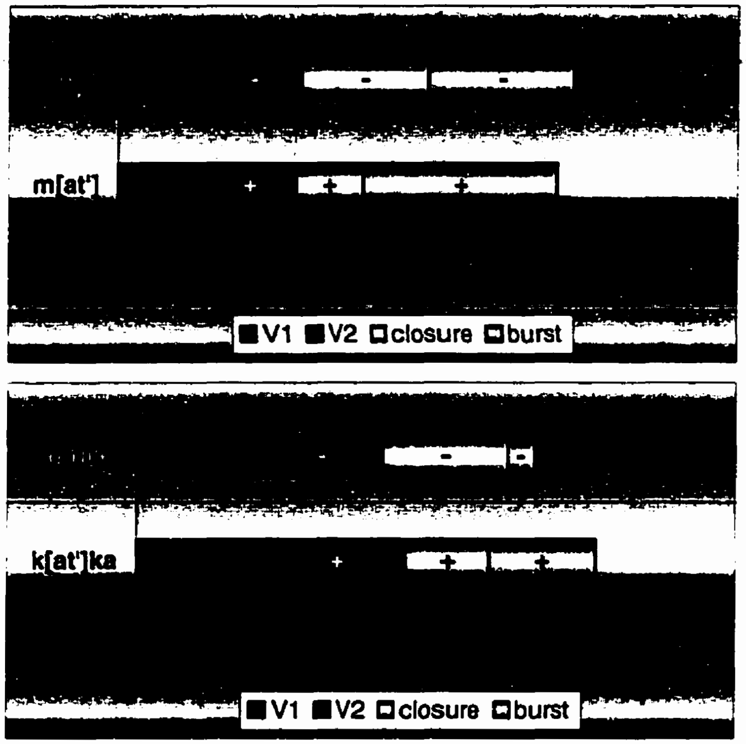


Figure 4.2. Minimal pairs (Speaker 1) with the vowel and final/preconsonantal consonant divided into four components, as used in Experiment 3.
"+" stands for "palatalized" and "-" is "plain".

Table 4.4. Identification scores of /t/ and /t'/ (.00 = /t/, 1.00 = /t'/) and response time (ms) based on various cue sets in final position. Means and standard deviations (in parentheses).

	Cues				*Identification*				*Response Time (ms)*	
	V1	*V2*	*Closure*	*Burst*	*Rate*		*%*			
							mat	*mat'*		
1	pal.	pal.	pal.	pal.	1.00	(.00)	0	100	541	(237)
2	pal.	pal.	*non*	pal.	1.00	(.00)	0	100	480	(128)
3	*non*	pal.	pal.	pal.	1.00	(.00)	0	100	445	(125)
4	*non*	pal.	*non*	pal.	.96	(.14)	4	96	505	(316)
5	pal.	pal.	pal.	*non*	.50	(.43)	50	50	885	(302)
6	*non*	*non*	*non*	pal.	.37	(.38)	63	37	835	(381)
7	*non*	*non*	pal.	pal.	.33	(.44)	67	33	678	(265)
8	*non*	pal.	*non*	*non*	.33	(.20)	67	33	873	(304)
9	pal.	pal.	*non*	*non*	.29	(.33)	71	29	804	(320)
10	*non*	pal.	pal.	*non*	.25	(.20)	92	25	716	(236)
11	pal.	*non*	pal.	pal.	.21	(.40)	79	21	801	(302)
12	pal.	*non*	*non*	pal.	.17	(.39)	83	17	681	(259)
13	pal.	*non*	pal.	*non*	.04	(.14)	96	4	572	(188)
14	*non*	*non*	pal.	*non*	.04	(.14)	96	4	547	(146)
15	pal.	*non*	*non*	*non*	.00	(.00)	1.00	0	495	(136)
16	*non*	*non*	*non*	*non*	.00	(.00)	1.00	0	513	(280)

Table 4.5. Identification scores of /t/ and /t'/ (.00 = /t/, 1.00 = /t'/) and response time (ms) based on various cue sets in preconsonantal position. Means and standard deviations (in parentheses).

	Cues				*Identification*				*Response Time (ms)*	
	V1	*V2*	*Closure*	*Burst*	*Rate*		% *mat*	% *mat'*		
1	pal.	pal.	pal.	pal.	1.00	(.00)	0	100	594	(337)
2	pal.	pal.	*non*	pal.	1.00	(.00)	0	100	594	(224)
3	*non*	pal.	pal.	pal.	1.00	(.00)	0	100	608	(178)
4	*non*	pal.	*non*	pal.	1.00	(.00)	0	100	575	(169)
5	*non*	*non*	*non*	pal.	0.96	(.14)	4	96	761	(378)
6	pal.	*non*	pal.	pal.	0.96	(.14)	4	96	853	(238)
7	*non*	*non*	pal.	pal.	0.92	(.19)	8	92	819	(420)
8	pal.	*non*	*non*	pal.	0.92	(.29)	8	92	729	(272)
9	pal.	pal.	pal.	*non*	0.58	(.42)	42	58	907	(340)
10	*non*	pal.	*non*	*non*	0.58	(.42)	42	58	736	(337)
11	pal.	pal.	*non*	*non*	0.54	(.50)	46	54	847	(275)
12	*non*	pal.	pal.	*non*	0.46	(.45)	54	46	792	(279)
13	pal.	*non*	pal.	*non*	.00	(.00)	100	0	561	(288)
14	*non*	*non*	pal.	*non*	.00	(.00)	100	0	617	(264)
15	pal.	*non*	*non*	*non*	.00	(.00)	100	0	554	(191)
16	*non*	*non*	*non*	*non*	.00	(.00)	100	0	646	(237)

The ANOVA results (Table 4.6) showed the significance of the factors of Burst [$F(1,11) = 245.49$; $p < .001$], V2 (transition) [$F(1,11) = 134.781$; $p < .001$] and Position [$F(1,11) = 19.161$; $p = .001$]. Note that Burst was the most important factor among these three (based on F values). Neither V1 ($p = 1.00$) nor Closure ($p = .678$) were significant.

Significant interactions between some of the factors were found. The most significant one is between Position, V2, and Burst [$F(1,11) = 25.07$; $p < .001$]. The interaction line plot (Figure 4.3) suggests that the two cue factors behave differently depending on the position. It is not clear how to explain the interaction between V1, V2, Closure, and Burst).

With respect to response time we found that only the interaction of the second half of the vowel (V2) and burst was significant (Table 4.7). As we see from Tables 4.4 and 4.5, RT values were relatively low when both V2 and burst had the same feature (plain or palatalized) and they were higher when the two major components were different with respect to palatalization.

Table 4.6 Identification of the /t/-/t'/ contrast based on various cue combinations

Source of Variance	*DF*	*F*	*p*
Position	1, 11	**19.16**	**.001**
V1	1, 11	.00	1.000
V2	1, 11	**134.78**	**.000**
Closure	1, 11	.18	.678
Burst	1, 11	**245.49**	**.000**
Position * V1	1, 11	.30	.594
Position *V2	1, 11	**15.70**	**.002**
Position * Closure	1, 11	1.60	.232
Position * Burst	1, 11	**14.08**	**.003**
V1 * V2	1, 11	**18.53**	**.001**
V1 * Closure	1, 11	4.31	.062
V1 * Burst	1, 11	**13.16**	**.004**
V2 * Closure	1, 11	.00	1.000
V2 * Burst	1, 11	.10	.760
Closure * Burst	1, 11	.27	.615
Position * V1 * V2	1, 11	3.47	.089
Position * V1 * Closure	1, 11	.11	.747
Position * V1 * Burst	1, 11	2.44	.147
Position * V2 * Closure	1, 11	1.69	.220
Position * V2 * Burst	1, 11	**25.07**	**.000**
Position * Closure * Burst	1, 11	1.74	.215
V1 * V2 * Closure	1, 11	1.07	.324
V1 * V2 * Burst	1, 11	.08	.782
V1 * Closure * Burst	1, 11	1.44	.255
V2 * Closure * Burst	1, 11	.06	.812
Position * V1 * V2 * Closure	1, 11	.14	.712
Position * V1 * V2 * Burst	1, 11	.77	.400
Position * V1* Closure * Burst	1, 11	.79	.418
Position * V2 * Closure * Burst	1, 11	.17	.688
V1 * V2 * Closure * Burst	1, 11	**4.82**	**.050**
Position * V1 * V2 * Closure * Burst	1, 11	.51	.490

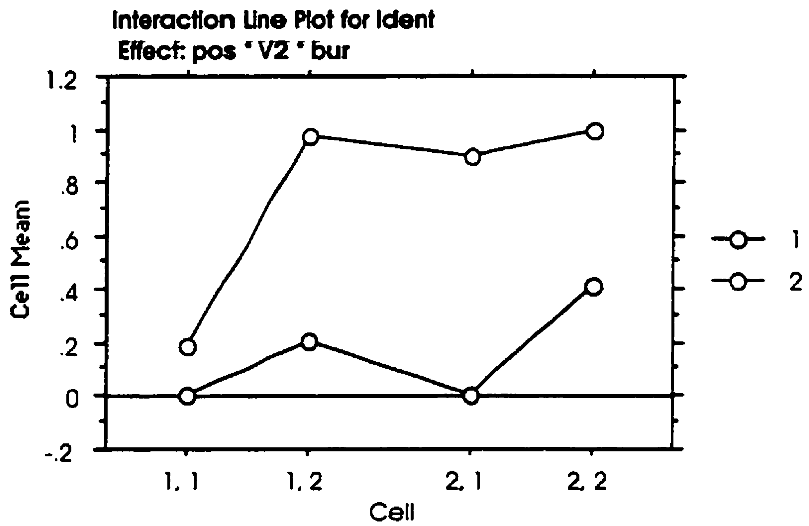


Figure 4.3. Interaction of factors Position, V2, and Burst. The trajectories represent the positions; V2 and burst values are given along X; Cell means are plotted along Y.

Table 4.7. Response time in identification of the /t/-/t'/ contrast based on various cue combinations. Only significant interactions are included.

Source of Variance	*DF*	*F*	*p*
Position	1, 11	2.66	.131
V1	1, 11	.36	.558
V2	1, 11	.35	.567
Closure	1, 11	1.05	.328
Burst	1, 11	1.52	.244
V2 * Burst	1, 11	**37.51**	**.000**

4.4.3 Discussion

The patterns of identification in the final and preconsonantal environments are somewhat different and will be discussed separately. Figure 4.4 presents the stimuli for the final contrast and their categorization by listeners. The identification curves allow one to divide the stimuli into three groups (areas). In Area I (stimuli 1 to 4) all stimuli have both palatalized burst and palatalized VC transition (Table 4.8). They vary in the values of V1 and closure. This variation has no apparent effect on identification: the response is /t'/ almost 100 per cent of the time. The average response time for these tokens is 493 ms.

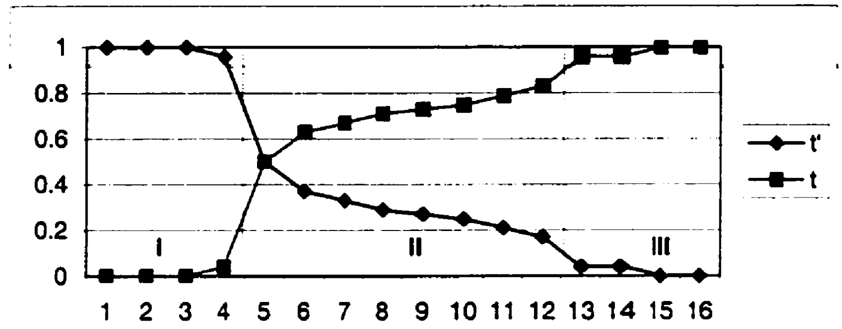


Figure 4.4. Identification of /t/ and /t'/ in final position: Rates of /t/ vs. /t'/ responses. (.00 = /t/, 1.00 = /t'/; random response at .5)

Table 4.8 Conditions and results in Experiment 5: final position

	Main conditions	*Results*	
Area I (stimuli 1-4)	*Burst* [pal] *Transition* [pal]	Consistent response: /t'/ Mean RT: 493 ms	*'easy'*
Area II (stimuli 5-12)	*Transition* [pal] --> [-pal] *Burst* [pal] or *Burst* [pal] --> [-pal] *Transition* [pal]	Random to less random response: /t/ Mean RT: 820 ms or Mean RT: 749 ms	*'hard'*
Area III (stimuli 13-16)	*Burst* [pal] --> [-pal] *Transition* [pal] --> [-pal]	Consistent response: /t/ Mean RT: 531 ms	*'easy'*

The common condition for the stimuli in Area II (stimuli 5-12) is a combination of two cues of opposite values: palatalized burst with a plain transition or plain burst with a palatalized transition. Other cues vary. As a result we observe a random response (listeners identify the stimuli either as /t'/ or as /t/) with a gradual favouring of /t/. RT is the highest at the beginning of the area (820 ms; almost twice as high as one in Area I) and somewhat lower later (749 ms). The stimuli of Area III have both plain burst and transition. This leads to a categorical identification of the signals as /t/. RT is now back to normal (531 ms).

The results show that the main cues to the contrast in the final position are transition and burst. The initial part of the vowel and closure do not seem to help the listener in making identification judgements. Transition is more important than burst in cueing the contrast.

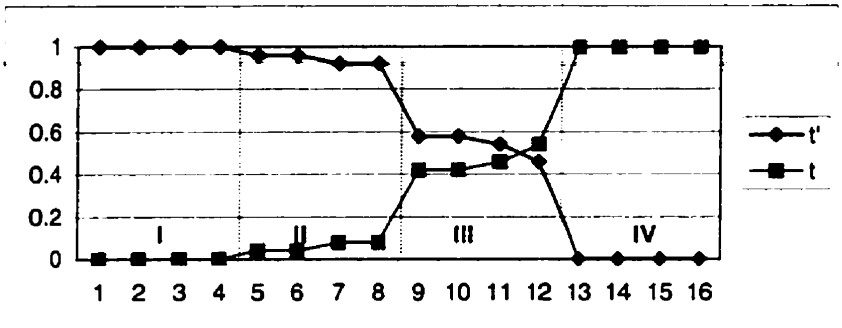


Figure 4.5. Identification of /t/ and /t'/ in preconsonantal position: cue manipulations. (.00 = /t/, 1.00 = /t'/; random response at .5)

Table 4.9 Conditions and results in Experiment 5, preconsonantal position

	Main conditions	*Results*	
Area I (stimuli 1-4)	*Burst* [pal] *Transition* [pal]	Consistent response: /t'/ Mean RT: 593 ms	*'easy'*
Area II (stimuli 5-8)	*Transition* [pal] --> [-pal] *Burst* [pal]	Less consistent response: /t'/ Mean RT: 791 ms	*'less easy'*
Area III (stimuli 9-12)	*Burst* [pal] --> [-pal] *Transition* [pal]	Random response Mean RT: 821 ms	*'hard'*
Area IV (stimuli 13-16)	*Burst* [pal] --> [-pal] *Transition* [pal] --> [-pal]	Consistent response: /t/ Mean RT: 595 ms	*'easy'*

Figure 4.5 and Table 4.9 describe the identification of stimuli for the preconsonantal contrast. Here we can divide the stimuli into four areas. As in the previous case, Area I shows a categorical identification of signals: the response is a constant /t'/. All stimuli here have both palatalized burst and VC transition. RT is fairly low, 593 ms. Changing the value of transition to palatalized, while keeping the burst palatalized (Area II), leads to less consistent, though still high, identification of /t'/. RT increases to 782 ms. Changing the value of the burst, while keeping the transition palatalized, gives us a dramatic effect: listeners respond at random. It takes much longer for the subjects to process the signal (815 ms). Cue sets with both plain burst and transition are consistently associated with /t/. RT goes down to 588 ms.

It is interesting to compare net effects of adding or removing a plain or palatalized burst or transition. In Table 4.10 I repeat selected tokens from Table 4.4 (final position). The common condition is that only one parameter, burst or transition, differs in value from all other components in the token. We can see that adding a plain burst to a non-palatalized token reduces its recognition as palatalized by 50% (the rate of .5 in Table 4.4). At the same time, adding a plain transition has a stronger effect of 79%. Note that the behaviour with respect to the palatalized burst or transition is not parallel to the plain cues.

Table 4.10. The net effect of adding bursts and transitions on identification in final position

Stimulus	*V1*	*V2*	*Closure*	*Burst*	*Effect, %*
5	pal.	pal.	pal.	***non***	50
11	pal.	***non***	pal.	pal.	79
6	*non*	*non*	*non*	**pal.**	37
8	*non*	**pal.**	*non*	*non*	33

Based on these results, the plain transition is more informative than plain burst (difference of 29%), while the palatalized transition conveys almost as much information as palatalized burst (difference of 4%). The plain vs. palatalized bursts and transitions are not equally informative: the plain components affect identification scores to a larger extent.

The situation in the preconsonantal position is quite different (Table 4.11). We find here that the plain burst is more informative than the plain transition (difference of 38%), which is a reverse of what we saw in Table 4.12. The palatalized burst is also a more robust cue here than the palatalized transition (difference of 38%). Both palatalized components affect the perception of the contrast to a greater extent than the plain ones. This is also the opposite of what we found for the final environment.

Table 4.11. The net effect of adding bursts and transitions on identification in preconsonantal position

Stimulus	*V1*	*V2*	*Closure*	*Burst*	*Effect, %*
5	pal.	pal.	pal.	***non***	42
11	pal.	***non***	pal.	pal.	4
6	*non*	*non*	*non*	**pal.**	96
8	*non*	**pal.**	*non*	*non*	58

In general, the findings suggest that transition conveys more information about the final stop and burst is the primary cue to the preconsonantal contrast. The responses show that the final position generally 'favours' the plain stop and the preconsonantal environment tends to 'prefer' the palatalized. It is not clear how these results can be straightforwardly explained by our acoustic findings. It is likely that the VC transition into a preconsonantal stop is affected by the following velar consonant. A further study should investigate whether this effect holds under other conditions and in other consonantal (e.g. following labial) and vocalic contexts.

To conclude, the effects of cues to palatalization present a complex pattern. In general, the burst is an important cue in identifying the contrast in *final* position; however, it is less crucial than the VC transition. The burst is more informative than the VC transition in the *preconsonantal* context. It takes much longer for a listener to identify a segment when the two major cues have contradictory values. The other components, closure and V1, do not seem to influence phonemic judgements.

4.5 Discussion and conclusion

In Chapters 3 and 4 I investigated the acoustic and perceptual correlates of the plain-palatalized contrast among final and preconsonantal coronal stops. I found that the stops are consistently released both finally and before another consonant. The plain and palatalized release bursts differ both spectrally and temporally. In terms of burst-to-closure ratio both consonants in final position can be described as affricates. The high frequency strident noise of coronal stops is accompanied by aspiration. In addition to burst, the two consonants differ in the duration of closure and in the spectral characteristics of VC transition. The VC transitions extend from about the mid point of the vowel, substantially affecting the vowel quality. While F2 rises before palatalized consonants, it somewhat lowers before plain stops. This may be due to the dental constriction of the stop or its secondary velarization.

The perception experiments determined that only VC transition and burst are perceptually relevant to the identification of the two phonemes. While there are significant durational differences between final and preconsonantal positions, they do not seem to affect the perception of the contrast, at least under the conditions of our experiments. The two environments, however, differ in the relevant importance of cues to the contrast: VC transition is more important for the final /t/ vs. /t'/ contrast, while burst is a major cue to the same contrast in preconsonantal position. The relative importance of plain and palatalized transitions and bursts also vary in the two environments.

These findings are important for our further experiments in Chapters 5 and 6, since it is the acoustic signal that is both a consequence of speaker's articulatory gestures and a source of information about the gestures for the listener.

Chapter 5. An articulatory study

5.0 Introduction

So far I have established that a study of the acoustics and perception of the plain-palatalized contrast provides insights into their distribution. In this chapter I switch my focus to articulation. A study of the articulation of this contrast in different environments is the key to testing our hypothesis, since according to our predictions (see section 1.3) neutralization is most likely to occur when certain gestures/features are poorly recovered as a result of articulatory factors, e.g. gestural reduction and overlap.

The chapter is organized as follows. First I review the previous work on articulation of Russian plain and palatalized stops. This, together with the results from Chapters 3 and 4, gives us an idea of what properties of plain and palatalized stops we should examine in our study. Then I discuss the setup for a series of articulatory experiments. The first experiment investigates the magnitude of the Tongue Body gesture at various points in time. The second experiment focuses on the primary articulators, Lips and Tongue Tip. Finally, the last study explores the interaction of primary gestures in clusters and the acoustic consequences of this process.

5.1 Previous work on Russian plain and palatalized consonants

I begin with a brief description of articulatory differences between the plain and palatalized coronal stops. Works on Russian phonetics (Avanesov 1984, Bolla 1981, Bondarko 1977, 1981, Halle 1959, Jones & Ward 1969, Kuznetsova 1965, 1969, Matusevich 1976, Skalozub 1963) show little disagreement about the primary articulation of plain and palatalized labials, as voiceless bilabial stops. There is a similar consensus about the production of the secondary articulation of /p'/, usually described as a secondary fronting and raising movement of tongue front (or tongue body, in the terminology used in the thesis) to the hard palate (cf. Ladefoged & Maddieson 1996: 364; Rogers 2000: 205).

/p/ is often described as having a secondary velar constriction (Matusevich 1976, Skalozub 1963), which is perceived as an off-glide before back vowels /u/, /ɔ/, and /i/ [ɨ] (Jones & Ward 1969: 79). Some researchers, however, disagree with this treatment of Russian 'plain' consonants. Halle 1959 (149-151) and Bolla 1981 note that non-palatalized consonants, including /t/, are better described as pharyngealized. According to these authors, non-palatalized consonants are characterized by narrowing of the pharynx, as opposed to its widening during the production of palatalized consonants. In this case, the retraction of the tongue body (velarization) is a by-product of narrowing the pharynx (pharyngealization). Ladefoged & Maddieson 1996: 361, however, argue that the term 'velarized' is appropriate only for the Russian non-palatalized lateral /l/.

The plain coronal /t/ is usually described as a stop articulated with the tip of the tongue or tongue blade against the upper teeth and alveolar ridge, i.e. as a denti-alveolar segment.[24] The active articulator is the tongue blade (Matusevich 1976: 133) or both the tongue blade and the tip of the tongue (laminal or apico-laminal) (Bolla 1981: 119-121; Jones & Ward 1969: 99). Just as for /p/, opinions about the secondary velar constriction of /t/ vary.

The palatalized coronal /t'/ has either almost the same constriction location as /t/ (Matusevich 1976) or it is slightly farther back in the alveolar (Bolla 1981: 119-121) or prepalatal area (at least for some speakers; Skalozub 1963: 27). The shape of the tongue, however, is more laminal, with the tip against the lower teeth. There is also a more significant lateral contact. The tongue body is fronted and raised to the hard palate, characteristic of a secondary palatal articulation. In addition to the main articulators, the lips are more spread during the palatalized /t'/ (Bolla 1981: 119-121). The articulatory differences between the four consonants are summarized in Table 5.1.

[24] The articulatory descriptions do not differentiate among stops in word-initial, internal, and final positions. In many cases they are likely to be based on word- or syllable-initial consonants (e.g. Bolla 1981).

Table 5.1. Articulation of plain and palatalized labial and coronal stops based on the literature

a.

Parameter	*/p/*	*/p'/*
Primary place	bilabial	bilabial
Secondary place	velarized/ pharyngealized	palatalized

b.

Parameter	*/t/*	*/t'/*
Primary place	denti-alveolar	(denti)-alveolar-(prepalatal)
Tongue shape	(apico)-laminal	laminal
Secondary place	velarized/ pharyngealized	palatalized

In my articulatory experiments I examine the properties that differentiate the four stops under consideration, focusing on their differences with respect to palatalization and variability across environments.

5.2 Experimental setup

5.2.0 Introduction

The goal of this section is to outline the methodology, procedure, and analysis involved in the series of articulatory experiments.

5.2.1 Procedure

Articulatory movement data were collected from three native speakers of Russian[25] using EMMA, the Haskins Electromagnetic Midsagittal Articulometer (Perkell et al. 1992). The articulometer has a set of small transducer coils, or receivers, that can be attached to various

[25] The subjects were a female from Moscow (Subject 1), a female from Perm' (Subject 2), and a male from Perm', the author of the thesis (Subject 3). Subject 1 has lived outside of Russia for 3 years. Subject 2 arrived in the U.S. a week prior to the experiment. Subject 3 had spent about 4 years outside Russia by the time of the experiment. All the speakers spoke Russian on a daily basis. The accent differences between the subjects were not considered to be of major importance for the study, since the two varieties of Standard Russian are noted to be very similar in the phonetic characteristics of consonants and deviate from each other mostly in the quality of unstressed vowels (Erofeeva 1997). The acoustic study, presented in Chapter 3 of this thesis, also identifies some burst (duration and spectrum) differences between the two accents. It should be mentioned that, unlike the Standard variety spoken in Perm', the older speakers of some rural Northern Russian dialects tend to neutralize final plain and palatalized labials and to have a pre-palatal realization of /t'/ (Shtern & Erofeeva 1998).

places of articulation. When a magnetic field is created, the movements of these coils, and thus the movements of articulators, are transmitted as a set of voltages that can be converted to distance. This method allows one to investigate the dynamics of the articulators, i.e. their movements in space (magnitude) and time (phasing or timing) (Figures 5.1 and 5.2).

In the current experiment the receivers were placed at the articulators of upper lip (UL), lower lip (LL), lower incisors, as an estimate of jaw movement, and several points on the tongue: tongue tip (TT), tongue body 1 (TB1), tongue body 2 (TB2), tongue dorsum (TD) (Figure 5.1). Estimates for the subjects' occlusal plane were also obtained as references to which the data could be rotated. The movement data were collected at a sampling rate of 500 Hz and the acoustic data at 20 kHz, using the real-time input software Maggie (Tiede et al. 1999). Further, the kinematic data were converted from voltage to distance, calibrated, and corrected for head movement and for any possible shifts of the receivers relative to the transmitters.

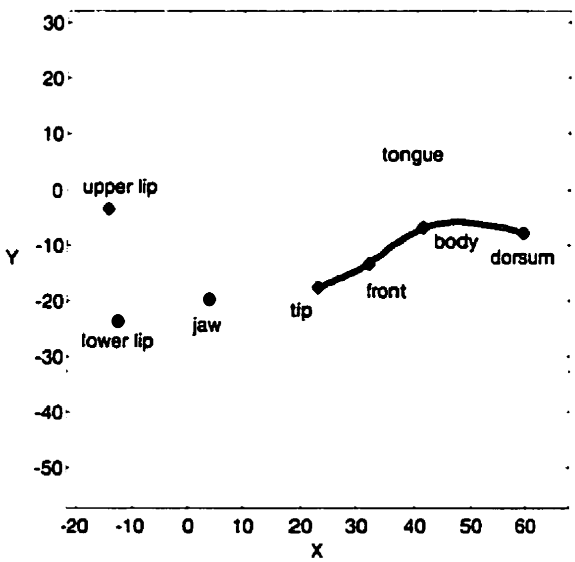


Figure 5.1 Places of articulators involved in the study.
The static view (sagittal section) during the utterance *eto pjapy opjat'*, Subject 3.

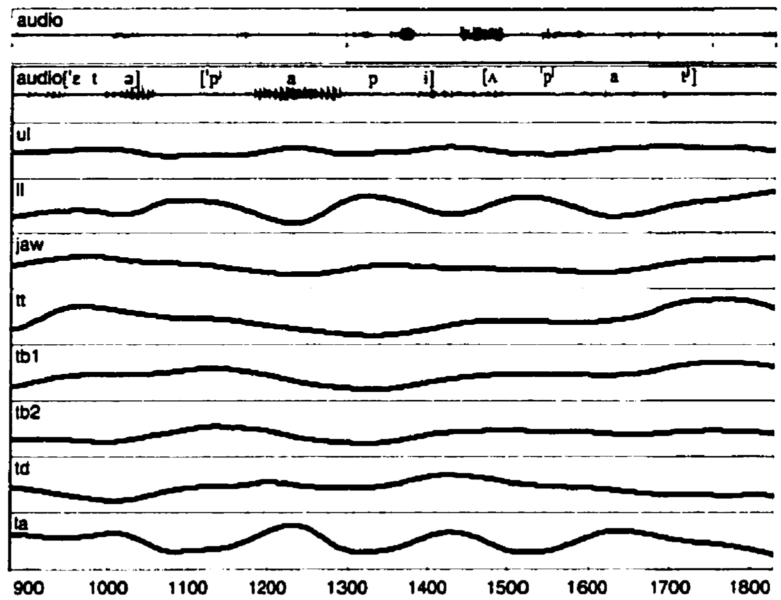


Figure 5.2. Trajectories of articulators during the utterance *eto pjapy opjat'*, Subject 3

There were some minor technical problems during the experimental sessions. A malfunction with a receiver coil on Tongue Dorsum after the second session of Subject 3 resulted in a defective set of data for this articulatory structure. The dorsals, however, were not the focus of the current experiment. There was also a problem with the Lower Lip receiver for Subject 1. The data for it was discarded and only the Upper Lip results were considered in the investigation of lip movements. Comparisons in terms of Upper Lip movement were made across the subjects to ensure the consistency of measurements.

5.2.2 Materials

The test phrases consisted of nonsense utterances of the type *ta(*C_1*)#(*C_2*)apy*, where C_1 and C_2 were any of the four stops, /p/, /p'/ /t/, and /t'/, as single consonants and in clusters (Table 5.2). The choice of nonsense stimuli was motivated by the purposes of the study. It allowed for all possible consonant combinations in the various identical environments. This would be difficult, if not impossible, to achieve with real words. Having word boundaries was important for testing our hypothesis, since not all clusters are phonotactically possible within Russian words (e.g. *p't or *t't) . And finally, using nonsense words allowed us to test perception both by native and non-native listeners, with no (or minimal) interference of lexical factors.

Table 5.2 Stimuli for the articulatory experiment

Type	*Environment*	*p*	*p'*	*t*	*t'*
Single	a#_a	ta papy	ta p'apy	ta tapy	ta t'apy
	a_#a	tap apy	tap' apy	tat apy	tat' apy
Clusters	a_#t	tap tapy	tap' tapy	tat tapy	tat' tapy
	a_#p	tap papy	tap' papy	tat papy	tat' papy
	a_#t'	tap t'apy	tap' t'apy	tat t'apy	tat' t'apy
	a_#p'	tap p'apy	tap' p'apy	tat p'apy	tat' p'apy

All of the of 36 stimuli were embedded in a frame sequence: *Eto _____ opját'* 'This is ____ again'.[26] The stress pattern was controlled for: both test words of the utterance were expected to have primary stress. *e.g. tá[t#p]ápy.*[27] While the syllabification of stops in clusters (C_1 and C_2 were hetero-syllabic, $C_1.C_2$) and of initial single stops (V.CV) was fairly straightforward, the syllable boundary is more difficult to determine with single final consonants (see Bondarko 1977: 126-129 on various views on syllabification in Russian).

Additional stimuli (the same for Subjects 1 and 2 but different for Subject 3) included nonsense utterances with other plain and palatalized consonants (e.g. /m/, /m'/, /n/, /n'/, /s/, /s'/, /j/) in a few selected environments and real utterances with cross-word and internal sequences with plain and palatalized stops (e.g. brá[t#p]ádaja 'brother falling' or o[tp]ál '(it) fell off'). The latter were used primarily in a separate study of gestural overlap (Kochetov & Goldstein 2001; see section 5.5.2.2).[28]

The stimuli were grouped in two blocks by type: nonsense and real. They were randomized and presented in Cyrillic on the screen of a computer using PowerPoint. The blocks of randomized and nonsense stimuli were alternated with blocks of real words (some of which were

[26] The stress is on the first syllable of the word *eto* and the last syllable of *opjat'*. The unstressed vowel /o/ in both cases is realized as [ʌ] (or as [ə] in faster speech).
[27] The results showed that the second word (1st syllable) had a higher level of stress compared to the first word.

used for another experiment). There were 5 repetitions. At the end of the experiment Subjects 1 and 2 were presented with a block of real words at a faster rate.

Due to changes during the course of the experiment, there was a minor difference in the stimuli (for the current study) presented to Subjects 1 and 2, compared to Subject 3. While the latter (tested at an earlier stage) was presented with the utterances *papy*, *p'apy*, *tapy*, and *t'apy* for initial single consonants, the former were given *ta papy*, *ta p'apy*, *ta tapy*, and *ta t'apy*, to keep the same number of syllables in the stimuli.

5.2.3 Analysis

In this section I report the data for plain and palatalized labial and coronal stops in nonsense phrases only. Some reference will be made to other results. For this experiment a total of 180 (36 phrases × 5 repetitions) articulatory tokens per subject were analyzed.

The articulatory and acoustic information for each token was stored in a separate file to be analyzed using the Mavis toolbox (Tiede et al. 1999) for MATLAB. To obtain more accurate characteristics of the lip movement (Lip Aperture), a tangential velocity, combining the data of both X and Y dimensions, was derived (Subjects 2 and 3; there was no valid Lower Lip movement for Subject 1).

Figure 5.3 presents a sample view of selected trajectories during the articulation of the sequence [ʌpʲa] (in *eto p'apy*).[29] The articulatory movements (the two lower images) are presented both in time (horizontal dimension, ms) and in space (vertical dimension, mm). The images above the gestures (audio) show the acoustic consequences of these movements, the segment corresponding to the gestures (lower) and the signal of the whole token (upper).

[28] There were a total of 65 nonsense and 93 real utterances for Subject 3 and 65 nonsense and 54 real utterances for Subjects 1 and 2. In addition, Subjects 2 and 3 were tested at a fast rate (15 real utterances).
[29] Note that the unstressed /o/ in *eto* is realized as [ʌ] or [ə].

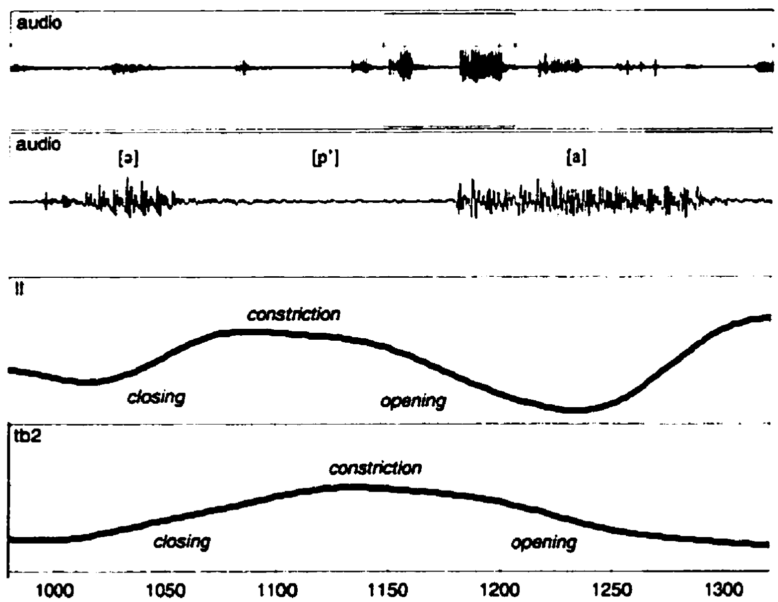


Figure 5.3 Dynamic view: Gestures in the articulation of initial [p'] in *eto p'apy*;
ll = lower lip, tb2 = tongue body; Subject 3

Looking at the trajectories, there are two gestures involved in the articulation of /p'/: Lips (represented here as the vertical movement of Lower Lip, LL) and Tongue Body (represented here as the vertical movement of tongue at point TB2). We can also identify the major components of these gestures: the closing movement (onset), constriction (plateau, or target) and the opening movement (offset, or release). Note that the constrictions of the two gestures are not simultaneous: the LL gesture precedes the Tongue Body gesture. The starting points of these components (the closing movement, constriction, and the opening movement) are further referred to as articulatory landmarks. Note also their relation to the acoustic signal (two upper images). The articulatory constriction roughly corresponds to (or rather results in) the acoustic closure, a period of silence of the stop. The release of the constriction results in an acoustic burst (not visible in this signal). Closing and opening gestural movements are roughly simultaneous with the preceding and following vowels. The information about these movements is available in the acoustic VC and CV transitions. I refer to the onset of acoustic closure, its offset (release), and release burst as acoustic landmarks.

We can also see the static view of the articulators at any given point in time. For instance, Figure 5.4 presents the sagittal view at the end of constriction of the Lower Lip gesture (about 1150, see Figure 5.3). The view has both vertical (y) and horizontal (x) dimensions. Notice that at this particular point in time the Tongue Body gesture is at its peak. We can measure its magnitude: it is at 30.74 mm on the horizontal scale and 1.44 mm on the vertical dimension. The choice of the initial point of the scale is arbitrary from subject to subject. Such displays allow us to compare the position of Tongue Body for the same consonant in various environments and between consonants in the same environments. The same can be done for any other gesture.

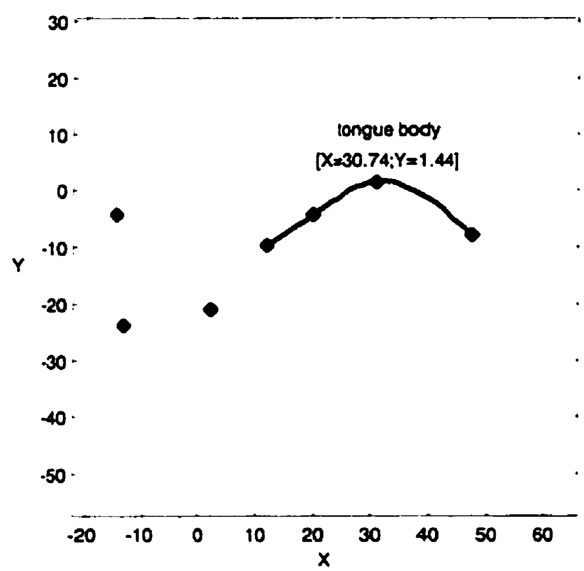


Figure 5.4 Static view: Position of articulators of initial [p'] in *eto p'apy* (at the end of ll constriction); Subject 3

In order to identify the landmarks of gestures (i.e. the crucial points: the beginning of closing movement, achievement of the target and its release), velocity for each articulator was obtained at thresholds of 20% of each of the examined articulators: Lip Aperture, Tongue Tip, and Tongue Body (i.e. the points on a trajectory where velocity exceeded threshold as movements). This was crucial in obtaining temporal information about the gestures and their phasing.

Figure 5.5 presents the gestures of Tongue Tip (a raising movement) and Lip Aperture (combined Lower and Upper Lips; a lowering movement) during the articulation of the cluster [tp] (in *tat papy*). Notice that the targets of two gestures (the shaded boxes) are timed far enough from each other to allow for a 50 ms lag between them. These measurements will also be important in our discussion (section 5.5.2).

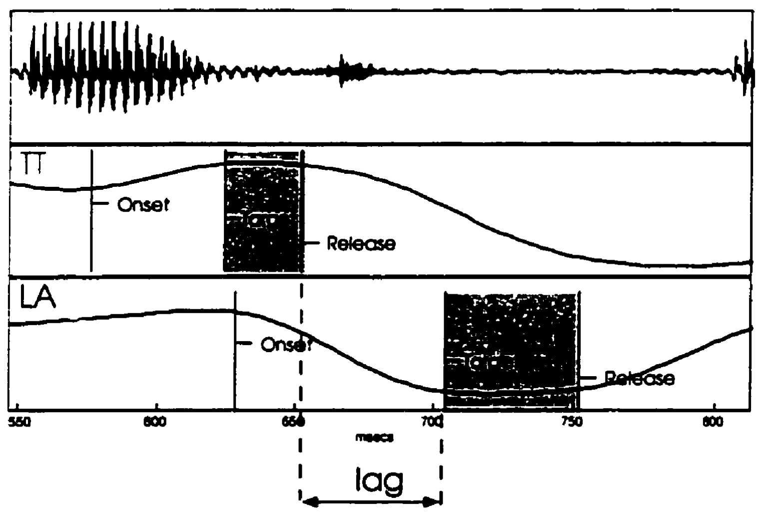


Figure 5.5. Timing of gestures of Tongue Tip and Lip Aperture in the cluster [tp] of tat papy and their landmarks (inset, target, and release); Subject 2

The investigation of the magnitude of gestures determines peaks and valleys of a particular movement (based on maximum and minimum values during the selected interval). The corresponding acoustic signal was also used in the analysis as a point of reference for measurements of Tongue Body magnitude.

The results obtained during the articulatory study were analyzed statistically using separate repeated analyses of variance (ANOVAs) with each of the measurements as the dependent variable. The speakers were analyzed separately, partly due to a relatively small sample size (3 speakers), and also due to some differences in measurements as a result of technical problems (see section 5.4: the Lower Lip trajectory was used for Subject 1, while Lip Aperture was evaluated for Subjects 2 and 3). Tukey HSD post-hoc tests were performed in each case in order to test significant interactions of factors. The analyses were performed with the package *Statistica*. The details of design will be presented separately for each variable.

5.3 Tongue Body magnitude: plain vs. palatalized

5.3.0 Introduction

The goal of the experiment presented in this section is to determine the articulatory difference between plain and palatalized consonants in terms of their secondary articulator, Tongue Body (TB). The first question of particular interest is: Does the Tongue Body movement differentiate plain and palatalized stops, and if so, how? The answer to this question is, overall, positive. Following this, another question arises. Are there any differences with respect to place of articulation and environment? Recall from section 5.2.3 that Tongue Body movement is determined by examining the magnitude (vertical and horizontal) of the movement of the tongue. See section 5.2 for the discussion of the experimental design.

5.3.1 Analysis

Measurements of Tongue Body magnitude (vertical and horizontal displacement) were made at four points in time, 30 ms prior to the acoustic closure of the stop (Point VC1), at the end of the first vowel and the onset of the closure (Point VC2), at the beginning of the second vowel (Point CV1), and 30 ms after this last point (Point CV2) (see section 5.3.2 and Figure 5.3 for details on acoustic landmarks).

The choice of acoustic landmarks (major spectral discontinuities, e.g. the onset and offset of a stop closure) rather than articulatory landmarks (major gestural discontinuities, e.g. onset or offset of gesture constriction) was motivated by the need for consistency in measurements of Tongue Body for palatalized and plain consonants as well as consistency in comparison of subjects. While the trajectory peaks or valleys can easily be determined for palatalized consonants, it is not easy to identify these for the plain consonants, which often show little or no deviation (particularly in clusters) and vary with respect to place of articulation, as well as across speakers. In addition, the choice of the acoustic landmarks is reasonable from the perceptual point of view: the acoustic information during the examined periods of the utterances is available to the listener through VC and CV transitions, while the trajectory during the closure is not readily available (except for the burst), and thus has to be deduced by the listener based on the transitions. In my discussion, however, I will refer to the both articulatory and acoustic landmarks.

To examine the parameters of interest I ran a number of separate ANOVAs. First I tested single consonants and then the coda consonant in clusters. Each test for *single* consonants involved three between-item factors, Place (labial and coronal), Palatalization (*Pal*: plain and palatalized), and Environment (*Env:* initial and final) and one dependent variable (TBy magnitude at points VC1, VC2, CV1, and CV2; TBx magnitude at the same points). Each test for the coda consonants in *clusters* involved four between-item factors, Place (labial and coronal), Palatalization (*Pal;* plain and palatalized), Environment Place (*EnvPlace*; labial and coronal), and Environment Palatalization (*EnvPal;* plain and palatalized) and one dependent variable (TBy magnitude at points VC1 and VC2; TBx magnitude at the same points). Each time Tukey HSD post-hoc tests were performed to investigate significant interactions.

5.3.2 Results: Single consonants

5.3.2.1 Overall results

In this section I give a general overview of the results. The results by major groups of stimuli as well as their discussion are presented in more detail in later sections. Tables 5.3-5.5 present means and standard deviations for Tongue Body magnitude, combined by dimension (vertical, TBy and horizontal, TBx) and by subject (based on the average per 5 tokens per subject). The consonants of interest are given in bold.

For TBy higher values represent higher position of Tongue Body in the oral cavity. For TBx lower values correspond to a more front position of Tongue Body in the oral cavity. Note that the absolute values cannot be compared across speakers, since they are based on arbitrary points of reference. In the case of the utterance *ta p'apy* as pronounced by Subject 1 (Table 5.3), at time VC1 the Tongue Body had a magnitude of -3.58 (.82) mm (horizontal displacement). The first number here stands for the mean value and the number in parentheses is the standard deviation. Later, at point VC2, the Tongue Body was raised to -.40 (.94) mm, and then further up to .84 (.65) mm at CV1. After that (at CV2) the Tongue Body gesture was lowered to -1.84 (.60) mm. The relevant overall results are discussed in section 5.3.2.6.

Table 5.3 Tongue Body magnitude at four points: Subject 1

a. Vertical displacement, TBy

	VC1		VC2		CV1		CV2	
	mean	SD	mean	SD	mean	SD	mean	SD
ta papy	-5.84	.82	-5.80	.64	-6.24	.73	-6.02	.76
ta p'apy	-3.58	1.23	-.40	.94	.84	.65	-1.84	.60
ta tapy	-5.10	.97	-1.86	.91	-2.94	.99	-5.50	.95
ta t'apy	-2.08	1.08	2.56	.99	.90	1.98	-2.00	2.13
tap apy	-5.00	1.76	-4.90	1.40	-6.02	1.62	-5.72	1.52
tap' apy	-4.28	.84	-.78	.91	-1.24	2.23	-4.02	1.96
tat apy	-4.92	.74	-1.50	.66	-2.34	.80	-4.82	1.21
tat' apy	-1.78	1.29	2.52	1.16	.96	1.42	-2.06	1.60

b. Horizontal displacement, TBy

	VC1		VC2		CV1		CV2	
	mean	SD	mean	SD	mean	SD	mean	SD
ta papy	29.42	1.07	31.66	1.85	33.96	.31	34.16	.58
ta p'apy	27.00	1.20	25.22	.97	25.16	1.00	28.28	1.39
ta tapy	28.02	.88	26.48	.87	26.54	.72	29.20	.87
ta t'apy	27.40	.73	25.80	.79	26.26	.55	29.52	1.02
tap apy	30.56	1.86	31.64	1.90	34.78	.33	34.94	.46
tap' apy	26.84	.94	25.48	.97	27.46	3.93	29.86	2.67
tat apy	28.68	.46	27.50	.92	27.48	1.01	30.24	.89
tat' apy	28.38	.88	26.80	.92	26.80	.64	30.02	.78

Table 5.4 Tongue Body magnitude at four points: Subject 2

a. Vertical displacement, TBy

	VC1		VC2		CV1		CV2	
	mean	SD	mean	SD	mean	SD	mean	SD
ta papy	.90	1.12	1.82	.79	1.06	1.27	.80	1.42
ta p'apy	2.02	.87	5.84	.94	8.76	.91	4.24	.69
ta tapy	.34	1.26	-2.24	.90	.96	.44	1.92	.53
ta t'apy	2.00	.82	7.90	1.05	9.04	.72	4.28	.48
tap apy	1.36	1.78	1.98	.59	1.02	.95	.68	.62
tap' apy	2.28	1.02	5.58	2.52	6.72	2.97	3.18	1.93
tat apy	-.40	1.49	-1.92	1.52	1.26	.48	2.12	1.92
tat' apy	2.80	1.96	8.70	1.10	8.60	.50	4.22	.84

b. Horizontal displacement, TBy

	VC1		VC2		CV1		CV2	
	mean	SD	mean	SD	mean	SD	mean	SD
ta papy	38.30	.91	42.40	.86	45.14	1.56	47.40	4.53
ta p'apy	37.86	.96	35.76	1.37	34.32	1.00	39.84	.61
ta tapy	37.64	.95	36.30	.89	38.88	.41	42.30	.63
ta t'apy	37.72	1.26	35.76	.70	37.08	.61	42.02	.54
tap apy	38.52	1.38	41.72	1.09	45.22	1.53	45.24	1.20
tap' apy	37.22	1.78	34.02	.91	36.00	4.21	41.58	2.17
tat apy	37.82	1.38	36.64	1.29	39.34	.75	42.14	1.82
tat' apy	37.88	2.18	35.80	1.15	37.98	.65	43.38	.87

Table 5.5 Tongue Body magnitude at four points: Subject 3

a. Vertical displacement, TBy

	VC1		VC2		CV1		CV2	
	mean	SD	mean	SD	mean	SD	mean	SD
ta papy	-4.34	1.42	-4.68	1.55	-6.34	1.36	-6.12	1.36
ta p'apy	-3.94	1.28	-1.62	1.25	1.64	.92	-1.80	.72
ta tapy	-4.40	1.10	-5.14	1.59	-5.58	2.26	-4.94	1.69
ta t'apy	-2.10	1.88	2.52	1.98	2.78	.97	-.92	.80
tap apy	-3.73	1.31	-3.30	1.52	-4.43	1.39	-4.48	1.41
tap' apy	-1.22	.83	.98	.46	-2.86	2.54	-4.02	2.49
tat apy	-4.98	1.33	-4.74	1.34	-3.82	1.31	-3.98	1.32
tat' apy	-1.38	1.47	2.48	1.40	.54	2.78	-1.98	2.04

b. Horizontal displacement, TBy

	VC1		VC2		CV1		CV2	
	mean	SD	mean	SD	mean	SD	mean	SD
ta papy	35.72	1.34	39.34	1.31	41.32	1.37	40.40	1.76
ta p'apy	35.16	.61	35.20	.93	29.64	.34	32.34	.46
ta tapy	36.68	1.35	35.82	1.51	36.18	1.81	38.44	1.92
ta t'apy	35.60	1.56	33.58	1.46	30.42	.69	34.22	1.04
tap apy	36.15	1.36	38.80	1.12	39.90	1.63	39.18	1.76
tap' apy	33.84	1.04	33.44	.59	37.94	1.90	38.98	2.25
tat apy	35.14	2.58	33.84	2.33	34.68	2.19	36.42	1.95
tat' apy	35.26	1.60	32.56	.92	32.76	1.94	36.68	1.40

The following sections (5.3.2.2-5.3.2.4) present statistical results for each subject and their summary (section 5.3.2.5). These are necessary to determine what factors and interactions are significant. The linguistically relevant discussion of the facts is given in section 5.3.2.6.

5.3.2.2 Results (statistics): Subject 1

Table 5.6 shows the ANOVA results for Tongue Body magnitude (TBy and TBx separately) for single consonants as pronounced by Subject 1. The significant results ($p < .05$) are given in bold. The main effects and significant interactions are further described point by point by displacement (TBy and TBx).

Table 5.6 Tongue Body magnitude: Subject 1 (statistics)

a. TBy

	VC1		VC2		CV1		CV2	
	F	*p*	*F*	*p*	*F*	*p*	*F*	*p*
Place	**11.29**	**.002**	**120.50**	**.000**	**26.49**	**.000**	3.13	.086
Pal	**40.58**	**.000**	**210.14**	**.000**	**112.01**	**.000**	**44.49**	**.000**
Env	.19	.669	.46	.503	.45	.509	.48	.494
Place × Pal	**4.91**	**.034**	.76	.390	**6.91**	**.013**	.04	.836
Place × Env	.06	.814	.03	.873	1.97	.170	1.89	.179
Pal × Env	.98	.330	1.84	.185	2.50	.123	3.13	.086
Place × Pal × Env	1.35	.256	.50	.483	.96	.334	.91	.346

b. TBx

	VC1		VC2		CV1		CV2	
	F	*p*	*F*	*p*	*F*	*p*	*F*	*p*
Place	.98	.331	**22.96**	**.000**	**53.91**	**.000**	**26.69**	**.000**
Pal	**27.09**	**.000**	**81.49**	**.000**	**77.12**	**.000**	**46.13**	**.000**
Env	3.73	.062	2.13	.154	**5.59**	**.024**	**5.95**	**.020**
Place × Pal	**14.81**	**.001**	**52.49**	**.000**	**60.76**	**.000**	**47.85**	**.000**
Place × Env	.24	.630	1.32	.259	.71	.405	.26	.612
Pal × Env	.52	.475	.03	.868	.31	.583	.03	.872
Place × Pal × Env	1.43	.241	.04	.848	.93	.341	.70	.408

5.3.2.2.1 Tongue Body: vertical displacement (TBy)

In this section I examine the statistical results for the vertical displacement of the Tongue Body at the different points of measurement taken point by point (see section 5.3.1 describing the points). Note that these and other results are discussed in section 5.3.2.6.

Point VC1

There was a main effect of Pal [F(1,20) = 40.58, p = .000]. Place was also significant [F(1,20) = 11.29, p = .002]. Env was not significant.

There was a significant Place × Pal interaction at point VC1 [F(1,20) = 4.91, p = .034]. This interaction was explored with the Tukey HSD test. The reason for the interaction was that the factor Place was significant for level two of the factor Pal, palatalized, /p'/ vs. /t'/ (p = .002), but not for level one of this factor, plain, /p/ vs. /t/ (.850). The factor Pal was significant for both levels of the factor Place (labial, /p/ vs. /p'/ (p = .030) and coronal, /t/ vs. /t'/ (p = .000)), but there was an asymmetry between levels labial (-5.42 vs. -3.93) and coronal (-5.01 vs. -1.93).

Point VC2

There was a main effect of Pal [F(1,20) = 210.14, p = .000]. Place was also significant [F(1,20) = 120.50, p = .000]. Env was not significant. There were no significant interactions.

Point CV1

There was a main effect of Pal [F(1,20) = 112.01, p = .000]. Place was also significant [F(1,20) = 26.49, p = .000]). Env was not significant.

There was a significant Place × Pal interaction at point CV1 [F(1,20) = 6.91, p = .013]. The reason for the interaction was that the factor Place was significant for level one of the factor Pal, plain, /p/ vs. /t/ (p = .000), but not for level two, palatalized, /p'/ vs. /t'/ (.301). The factor Pal was significant for both levels of the factor Place (labial, /p/ vs. /p'/ (p = .000) and coronal, /t/ vs. /t'/ (p = .000)).

Point CV2

There was a main effect of Pal [F(1,20) = 44.49, p = .000]). Place and Env were not significant. There were no significant interactions.

In sum, there was a main effect of Pal throughout the examined CV-VC period. Place was also significant, except for the last point, CV2. Env was not significant.

5.3.2.2.2 Tongue Body: horizontal displacement (TBx)

This section presents the statistical results for the horizontal displacement of the Tongue Body at the different points of measurement taken point by point. Note that these and other results are discussed in section 5.3.2.6.

Point VC1

There was a significant Place × Pal interaction [$F(1,20) = 14.81$, $p = .001$]. The reason for the interaction is that the factor Pal was significant for level one (labial, /p/ vs. /p'/) of the factor Place ($p = .000$) but not for level two (coronal, /t/ vs. /t'/) of this factor ($p = .773$). Also, the factor Place was significant for level one (plain, /p/ vs. /t/) of the factor Pal ($p = .009$) but not for level two (palatalized, /p'/ vs. /t'/) of this factor ($p = .201$).

Point VC2

There was a significant Place × Pal interaction [$F(1,20) = 52.49$, $p = .000$]. The reason for the interaction is that the factor Pal was significant for level one (labial, /p/ vs. /p'/) of the factor Place ($p = .000$) but not for level two (coronal, /t/ vs. /t'/) of this factor ($p = .594$). Also, the factor Place was significant for level one (plain /p/ vs. /t/) of the factor Pal ($p = .000$) but not for level two (palatalized, /p'/ vs. /t'/) of this factor ($p = .323$).

Point CV1

There was a significant Place × Pal interaction [$F(1,20) = 60.76$, $p = .000$]. The reason for the interaction is that the factor Pal was significant for level one (labial, /p/ vs. /p'/) of the factor Place ($p = .000$) but not for level two (coronal, /t/ vs. /t'/) of this factor ($p = .897$). Also, the factor Place was significant for level one (plain /p/ vs. /t/) of the factor Pal ($p = .000$) but not for level two (palatalized, /p'/ vs. /t'/) of this factor ($p = .989$).

Point CV2

There was a significant Place × Pal interaction [$F(1,20) = 47.85$, $p = .000$]. The reason for the interaction is that the factor Pal was significant for level one (labial, /p/ vs. /p'/) of the factor Place ($p = .000$) but not for level two (coronal, /t/ vs. /t'/) of this factor ($p = 1.000$). Also, the factor Place was significant for level one (plain /p/ vs. /t/) of the factor Pal (p = .000) but not for level two (palatalized, /p'/ vs. /t'/) of this factor (p = .608).

In sum, there was a main effect of Pal throughout the examined CV-VC period. Place was also significant, except for the first point, VC1. Env was significant at both CV points, CV1 and CV2.

5.3.2.3 Results (statistics): Subject 2

Table 5.7 shows the ANOVA results for Tongue Body magnitude (TBy and TBx separately) for single consonants as pronounced by Subject 2.

Table 5.7 Tongue Body magnitude: Subject 2 (statistics)

a. TBy

	VC1		VC2		CV1		CV2	
	F	*p*	*F*	*p*	*F*	*p*	*F*	*p*
Place	1.32	.259	2.83	.102	1.98	.169	**5.76**	**.02**
Pal	**19.00**	**.000**	**295.37**	**.000**	**310.29**	**.000**	**47.03**	**.00**
Env	.24	.626	.38	.541	1.84	.184	.47	.50
Place × Pal	3.17	.084	**63.32**	**.000**	1.52	.226	.95	.34
Place × Env	.17	.680	.55	.465	1.41	.244	.76	.39
Pal × Env	.72	.404	.00	.971	2.81	.104	.63	.44
Place × Pal × Env	1.21	.280	.30	.590	.59	.447	.20	.66

b. TBx

	VC1		VC2		CV1		CV2	
	F	*p*	*F*	*p*	*F*	*p*	*F*	*p*
Place	.22	.642	**49.65**	**.000**	**10.95**	**.002**	2.80	.104
Pal	.80	.378	**138.86**	**.000**	**107.60**	**.000**	**16.55**	**.000**
Env	.00	.965	2.34	.136	1.95	.173	.10	.759
Place × Pal	1.11	.301	**94.38**	**.000**	**56.96**	**.000**	**23.32**	**.000**
Place × Env	.18	.674	**4.41**	**.044**	.03	.859	.41	.525
Pal × Env	.24	.626	1.04	.316	.83	.369	**4.62**	**.039**
Place × Pal × Env	.22	.642	.33	.573	.27	.608	.89	.352

5.3.2.3.1 Tongue Body: vertical displacement (TBy)

In this section I examine the statistical results for the vertical displacement of the Tongue Body at different points of measurement taken point by point.

Point VC1

There was a main effect of Pal [F(1,20) = 19.00, p = .000]. Place and Env were not significant. There were no significant interactions.

Point VC2

There was a main effect of Pal [F(1,20) = 295.37, p = .000]. Place and Env were not significant.

There was a significant Place × Pal interaction [F(1,20) = 63.32, p = .000]. The factor Pal was significant for both level one (labial, /p/ vs. /p'/) of the factor Place (p = .000) and level two (coronal, /t/ vs. /t'/) of this factor (p = .000). Also, the factor Place was significant for both level one (plain /p/ vs. /t/) of the factor Pal (p = .000) and level two (palatalized, /p'/ vs. /t'/) of this factor (p = .000). The interaction can be explained by the asymmetry of the mean values.

Point CV1

There was a main effect of Pal [F(1,20) = 310.29, p = .000]. Place and Env were not significant. There were no significant interactions.

Point CV2

There was a main effect of Pal [F(1,20) = 47.03, p = .000]). Place was also significant [F(1,20) = 5.76, p = .02]. Env was not significant. There were no significant interactions.

In sum, there was a main effect of Pal throughout the examined CV-VC period. Place was significant only at the last point, CV2. Env was not significant.

5.3.2.3.2 Tongue Body: horizontal displacement (TBx)

In this section I examine the statistical results for the horizontal displacement of the Tongue Body at the different points of measurement taken point by point

Point VC1

None of the factors was significant. There were no significant interactions.

Point VC2

There was a main effect of Pal [F(1,20) = 138.86, p = .000]. Place was also significant [F(1,20) = 49.65, p = .000]. Env was not significant.

There was a significant Place × Pal interaction [F(1,20) = 94.38, p = .000]. The reason for the interaction is that the factor Pal was significant for level one (labial, /p/ vs. /p'/) of the factor Place (p = .000) but not for level two (coronal, /t/ vs. /t'/) of this factor (p = .471). Also, the factor Place was significant for level one (plain /p/ vs. /t/) of the factor Pal (p = .000) but not for level two (palatalized, /p'/ vs. /t'/) of this factor (p = .254).

There was a significant Place × Env interaction [F(1,20) = 4.41, p = .044]. The reason for the interaction is that the factor Env was significant for level one (labial, /p/ vs. /p'/) of the factor Place (p = .000) but not for level two (coronal, /t/ vs. /t'/) of this factor (p = .978). The factor Place was significant for both level one (initial, #labial vs. #coronal) of the factor Env (p = .000) and level two (final, labial# vs. coronal#) of this factor (p = .007).

Point CV1

There was a main effect of Pal [F(1,20) = 107.60, p = .000]. Place and Env were not significant.

There was a significant Place × Pal interaction [F(1,20) = 56.96, p = .000]. The reason for the interaction is that the factor Pal was significant for level one (labial, /p/ vs. /p'/) of the factor Place (p = .000) but not for level two (coronal, /t/ vs. /t'/) of this factor (p = .210). The factor Place was significant for both level one (plain /p/ vs. /t/) of the factor Pal (p = .000) and level two (palatalized, /p'/ vs. /t'/) of this factor (p = .026).

Point CV2

There was a main effect of Pal [F(1,20) = 16.55, p = .000]). Place was also significant [F(1,20) = 10.95, p = .002]. Env was not significant.

There was a significant Place × Pal interaction [F(1,20) = 23.32, p = .000]. The reason for the interaction is that the factor Pal was significant for level one (labial, /p/ vs. /p'/) of the factor Place (p = .000) but not for level two (coronal, /t/ vs. /t'/) of this factor (p = .949). Also, the factor Place was significant for level one (plain /p/ vs. /t/) of the factor Pal (p = .000) but not for level two (palatalized, /p'/ vs. /t'/) of this factor (p = .137).

There was a significant Pal × Env interaction [F(1,20) = 4.62, p = .039]. The reason for the interaction is that the factor Pal was significant for level one (initial, #C vs. #C') of the factor Env (p = .001) but not for level two (final, C# vs. C'#) of this factor (p = .535). The factor Env was significant neither for level two (palatalized, #C' vs. C'#) of the factor Pal (p = .569) nor level one (plain, #C vs. C#) of this factor (p = 321).

In sum, there was a main effect of Pal at all points, except for the first one, VC1. Place was also significant at two points, CV2 and VC2. Env was not significant.

5.3.2.4 Results (statistics): Subject 3

Table 5.8 shows the ANOVA results for Tongue Body magnitude (TBy and TBx separately) for single consonants as pronounced by Subject 3.

Table 5.8 Tongue Body magnitude: Subject 3 (statistics)

a. TBy

	VC1		VC2		CV1		CV2	
	F	*p*	*F*	*p*	*F*	*p*	*F*	*p*
Place	.05	.833	**4.20**	**.049**	**6.55**	**.015**	**5.30**	**.028**
Pal	**26.27**	**.000**	**148.21**	**.000**	**93.18**	**.000**	**29.24**	**.000**
Env	*4.09*	*.051*	**5.65**	**.024**	1.77	.193	.11	.738
Place × Pal	3.04	.091	**17.07**	**.000**	1.90	.178	.39	.537
Place × Env	3.46	.072	3.93	.056	.83	.368	.06	.813
Pal × Env	*3.93*	*.056*	.18	.672	**20.39**	**.000**	**8.69**	**.006**
Place × Pal × Env	.22	.643	.83	.370	1.10	.303	.85	.362

b. TBx

	VC1		VC2		CV1		CV2	
	F	*p*	*F*	*p*	*F*	*p*	*F*	*p*
Place	.88	.355	**40.58**	**.000**	**52.90**	**.000**	**5.98**	**.020**
Pal	*3.94*	*.056*	**57.05**	**.000**	**110.37**	**.000**	**33.86**	**.000**
Env	2.06	.161	**9.45**	**.004**	**14.47**	**.001**	**7.78**	**.009**
Place × Pal	.98	.330	**12.04**	**.002**	**8.63**	**.006**	**4.19**	**.049**
Place × Env	.26	.611	.17	.687	**8.86**	**.006**	**5.62**	**.024**
Pal × Env	.08	.777	.02	.881	**44.65**	**.000**	**34.59**	**.000**
Place × Pal × Env	2.34	.136	1.60	.215	**8.40**	**.007**	2.60	.117

5.3.2.4.1 Tongue Body: vertical displacement (TBy)

In this section I examine the statistical results for the vertical displacement of the Tongue Body at different points of measurement taken point by point.

Point VC1

There was a main effect of Pal [F(1,20) = 26.27, p = .000]. Env was slightly above the significance level ([F(1,20) = 3.94, p = .056]). Place was not significant. There were no significant interactions.

Point VC2

There was a main effect of Pal [F(1,20) = 148.21, p = .000]. Place [F(1,20) = 4.20, p = .049] and Env [F(1,20) = 5.65, p = .024] were also significant.

There was a significant Place × Pal interaction [F(1,20) = 17.07, p = .000]. The reason for the interaction is that the factor Place was significant for level two (palatalized, /p'/ vs. /t'/) of the factor Pal (p = .000) but not for level one (plain /p/ vs. /t/) of this factor (p = .466). The factor Pal was significant for both level one (labial, /p/ vs. /p'/) of the factor Place (p = .000) and level two (coronal, /t/ vs. /t'/) of this factor (p = .000).

Point CV1

There was a main effect of Pal [F(1,20) = 93.18, p = .000]. Place was also significant [F(1,20) = 6.55, p = .015]. Env was not significant.

There was a significant Pal × Env interaction [F(1,20) = 20.39, p = .000]. The factor Pal was significant for both level one (initial, #C vs. #C') of the factor Env (p = .000) and level two (final, C# vs. C'#) of this factor (p = .005). Also, the factor Env was significant for level two (palatalized, #C' vs. C'#) of the factor Pal (p = .005), but not for level one (plain, #C vs. C#) of this factor (p = .130).

Point CV2

There was a main effect of Pal [F(1,20) = 29.24, p = .000]). Place was also significant [F(1,20) = 5.30, p = .028]. Env was not significant.

There was a significant Pal × Env interaction [F(1,20) = 8.69, p = .006]. The reason for the interaction is that the factor Pal was significant for level one (initial, #C vs. #C') of the factor

Env (p = .000) but not for level two (final, C# vs. C'#) of this factor (p = .319). The factor Env was significant neither for level two (palatalized, #C' vs. C'#) of the factor Pal (p = .114) nor level one (plain, #C vs. C#) of this factor (p = .273).

In sum, there was a main effect of Pal throughout the examined CV-VC period. Place was always significant, except for the first point, VC1. Env was significant at Point VC2 and slightly above the significance level at point VC1.

5.3.2.4.2 Tongue Body: horizontal displacement (TBx)

In this section I examine the statistical results for the horizontal displacement of the Tongue Body at different points of measurement taken point by point.

Point VC1

None of the factors was significant. There were no significant interactions.

Point VC2

There were main effects of Pal, Place, and Env (Pal: [F(1,20) = 40.58, p = .000]; Place: [F(1,20) = 57.05, p = .000]; Env.: [F(1,20) = 9.45, p = .004]).

There was a significant Place × Pal interaction [F(1,20) = 12.04, p = .002]. The reason for the interaction is that the factor Pal was significant for level one (labial, /p/ vs. /p'/) of the factor Place (p = .000) but not for level two (coronal, /t/ vs. /t'/) of this factor (p = .057). Also, the factor Place was significant for level one (plain /p/ vs. /t/) of the factor Pal (p = .000) but not for level two (palatalized, /p'/ vs. /t'/) of this factor (p = .258).

Point CV1

There were main effects of Pal, Place, and Env [F(1,20) = 52.90, p = .000], [F(1,20) = 5.98, p = .020]; Place: [F(1,20) = 110.37, p = .000]; Env.: [F(1,20) = 14.47, p = .001]).

There was a significant Place × Pal interaction [F(1,20) = 8.63, p = .006]. The factor Place was significant for level one (plain /p/ vs. /t/) of the factor Pal (p = .000) and level two (palatalized, /p'/ vs. /t'/) of this factor (p = .022). The factor Pal was significant for both level one (labial, /p/ vs. /p'/) of the factor Place (p = .000) and level two (coronal, /t/ vs. /t'/) of this factor (p = .000). The reason for the interaction is the asymmetry between the means: plain labial /p/: 40.61, palatalized labial /p'/: 33.79, plain coronal /t/: 35.43, palatalized coronal /t'/: 31.59. the difference between /p/ and /p'/ (Pal) is greater than between /t/ and /t'/. The same holds for Place: a larger difference between /p/ vs. /t/ than between /p'/ and /t'/. The interaction can be explained by the asymmetry of the mean values (Means: /p/ 40.61, /p'/ 33.79, /t/ 35.43, /t'/ 31.59).

There was a significant Pal × Env interaction [$F(1,20) = 44.65$, $p = .000$]. The reason for the interaction is that the factor Pal was significant for level one (initial, #C vs. #C') of the factor Env (p = .000) but it was marginal for level two (final, C# vs. C'#) of this factor (p = .051). Also, the factor Env was significant for level two (palatalized, #C' vs. C'#) of the factor Pal (p = .000) but not for level one (plain, #C vs. C#) of this factor (p = .197).

There was a significant Place × Env interaction [$F(1,20) = 8.86$, $p = .006$]. The reason for the interaction is that the factor Env was significant for level one (labial, /p/ vs. /p'/) of the factor Place (p = .000) but not for level two (coronal, /t/ vs. /t'/) of this factor (p = .936). The factor Place was significant for both level one (initial) (p = .023) and level two (final) of the factor Env (p = .000) of this factor.

There was a significant Place × Pal × Env interaction [$F(1,20) = 8.40$, $p = .007$]. The reason for the interaction is that the factor Env was significant for some combinations of levels of Place and Pal (/p'/ (p = .001)), but not for others (/p/ (p = .851), /t/ (p = .813), /t'/ (p = .321)). The factor Pal was significant for some combinations of levels of Place and Env (#labial (p = .000), #coronal (p = .000)), but not for others (labial# (p = .541), coronal# (p = .566)). Also, the factor Place was significant for some combinations of levels of Pal and Env (#C (p = .001), C# (p = .000), C'# (p = .000)), but not for others (#C' (p = .994)).

Point CV2

There was a main effect of Pal, Place, and Env [F(1,20) = 5.98, p = .020]; Place: [F(1,20) = 33.86, p = .000]; Env.: [F(1,20) = 7.78, p = .009]).

There was a significant Place × Pal interaction [F(1,20) = 4.19, p = .049]. The reason for the interaction is that the factor Pal was significant for level one (labial, /p/ vs. /p'/) of the factor Place (p = .000) but not for level two (coronal, /t/ vs. /t'/) of this factor (p = .055). Also, the factor Place was significant for level one (plain /p/ vs. /t/) of the factor Pal (p = .016) but not for level two (palatalized, /p'/ vs. /t'/) of this factor (p = .992).

There was a significant Pal × Env interaction [F(1,20) = 34.59, p = .000]. The reason for the interaction is that the factor Pal was significant for level one (initial, #C vs. #C') of the factor Env (p = .000) but not for level two (final, C# vs. C'#) of this factor (p = 1.000). Also, the factor Env was significant for level two (palatalized, #C' vs. C'#) (p = .000) of the factor Pal but not for level one (plain, #C vs. C#) of this factor (p = .150).

There was a significant Place × Env interaction [F(1,20) = 5.62, p = .024]. The reason for the interaction is that the factor Place was significant for level two (final) of the factor Env (p = .009) but not for level one (initial) of this factor (p = 1.000). Also, the factor Env was significant for level one (labial, /p/ vs. /p'/) of the factor Place (p = .005) but not for level two (coronal, /t/ vs. /t'/) of this factor (p = . 991). In sum, there was a main effect of Pal, Place, and Env at all points, except for VC1.

5.3.2.5 Summary

In this section I summarize the statistical results for Tongue Body magnitude.

Tongue Body: vertical displacement (TBy)

Pal is significant throughout both VC and CV transitions for all three subjects. Place is more variable: it is significant either at the first three points, VC1, VC2, and CV1 (Subject 1), at the last three points, VC2, CV1, and CV2 (Subject 3), or at the last point, CV2 (Subject 3). Env is significant at VC2 (Subject 3), or not significant at all (Subjects 1 and 2).

Among the interactions, Place × Pal is significant at VC1 and CV2 (Subject 1), or VC2 (Subject 2 and Subject 3). The interaction Pal × Env is significant only for Subject 3 (at CV1 and CV2).

Tongue Body: horizontal displacement (TBx)

Pal is significant throughout both VC and CV transitions (Subject 1) or the last three points, VC1, CV1, CV2 (Subject 2 and Subject 3). Place is significant either at the last three points, VC1, CV1, and CV2 (Subject 1 and Subject 3), or at the stop onset and release, VC2 and CV1 (Subject 2). Env is significant at the three last points, VC2, CV1, and CV2 (Subject 3), at the CV transition, CV1 and CV2 (Subject 1), or is not significant at all (Subject 2).

Among the interactions, Place × Pal is significant throughout both VC and CV transitions (Subject 1) or at the three last points, VC2, CV1, and CV2 (Subject 2 and Subject 3). The interaction Place × Env is significant at the CV2 transition, CV1 and CV2 (Subject 3), at VC2 (Subject 2) or is not significant (Subject 1). The interaction Pal × Env is significant at the CV2 transition, CV1 and CV2 (Subject 3), at point CV2 (Subject 2) or is not significant (Subject 1). The interaction Place × Pal × Env is significant only for Subject 3 (at CV1).

Overall, Pal is highly significant for most of the transitions for all the speakers. It is followed by Place, and further, by Environment.

5.3.2.6 Discussion

In this section I discuss the articulatory (TB) differences between stops in terms of palatalization (plain and palatalized), place (labial and coronal), and (environment (initial and final).

5.3.2.6.1 Palatalization

Here I address the Tongue Body differences between plain and palatalized labials (/p/ vs. /p'/) and coronals (/t/ vs. /t'/), and compare these two contrasts with each other.

5.3.2.6.1.1 /p/ vs. /p'/

Figure 5.6 presents the Tongue Body trajectories of /p/ and /p'/ for the three speakers. The trajectories are based on the measurements at four points in time, discussed in the previous section: Point 1 (VC1), at 30 ms before the acoustic closure of the stop, Point 2 (VC2), or at the onset of the constriction, Point 3 (CV1), at the onset of the vowel, and Point 4 (CV2), at 30 ms after the onset of the vowel. The vertical displacement (TBy) is given in the left column and the horizontal displacement (TBx) is in the right column. For TBy the higher values represent higher position of Tongue Body in the oral cavity. For TBx the lower values correspond to a more front position of Tongue Body in the oral cavity. The dotted line represents the trajectory during the stop constriction (closure and burst).

The overall Tongue Body differences between /p/ and /p'/ at each point are given in Table 5.9. Non-significant differences (based on the corresponding main effects of Pal and significant Place × Pal interactions) are marked with an asterisk. The highest values per category are given in bold. Note that the differences between points in time were not evaluated statistically.

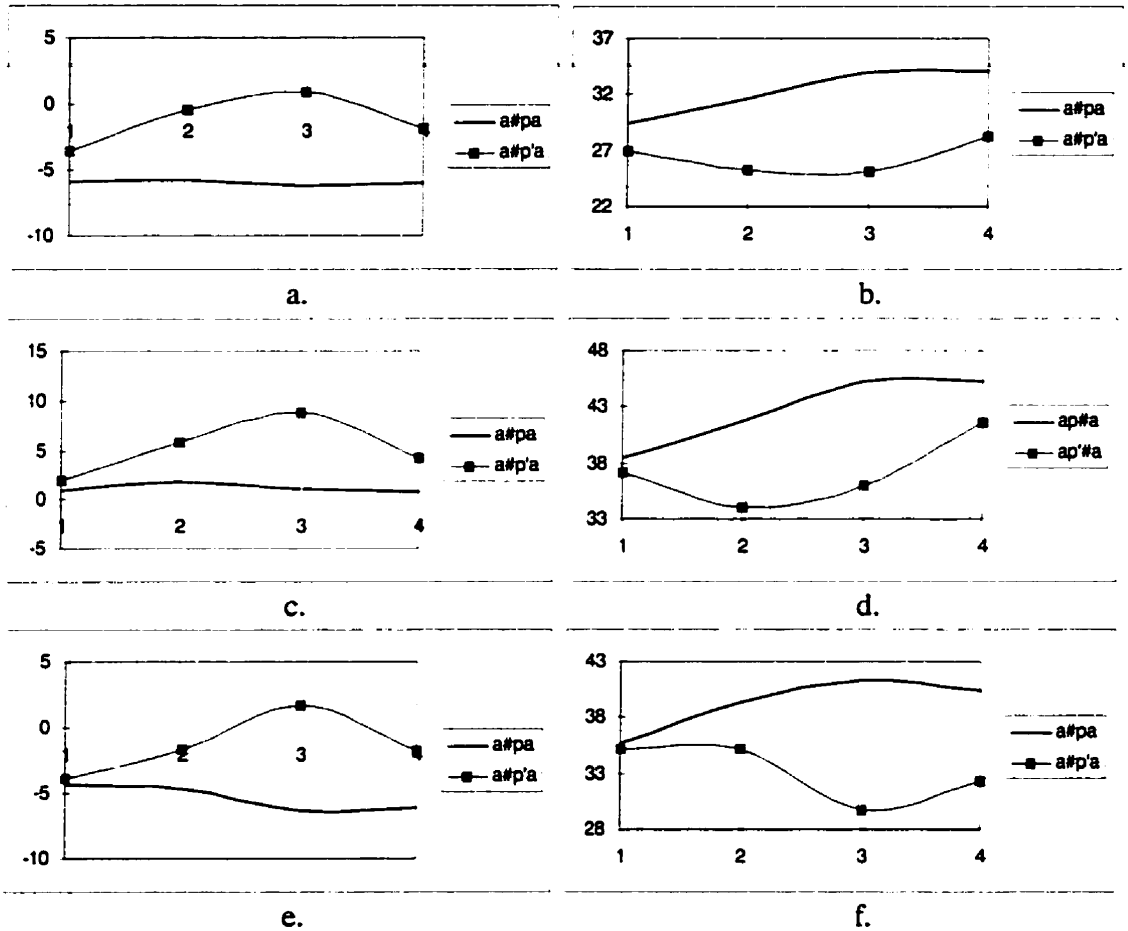


Figure 5.6 Movement of Tongue Body (vertical (ace) and horizontal (bdf)) during the articulation of /p/ and /p'/, based on means at four points in time for Subjects 1 (ab), 2 (cd), and 3 (ef).

Table 5.9 Mean differences between initial /p/ and /p'/

		VC1	VC2	CV1	CV2	Total VC	Total CV	Total	VC, %	CV, %
TBy	S1	2.26	5.4	**7.08**	4.18	7.66	**11.26**	*18.9*	*0.4*	***0.6***
	S2	1.12	4.02	**7.7**	3.44	5.14	**11.14**	*16.3*	*0.32*	***0.68***
	S3	0.4	3.06	**7.98**	4.32	3.46	**12.3**	*15.8*	*0.22*	***0.78***
TBx	S1	2.42	6.44	**8.8**	5.88	8.86	**14.68**	*23.5*	*0.38*	***0.62***
	S2	0.44*	6.64	**10.82**	7.56	7.08	**18.38**	*25.5*	*0.28*	***0.72***
	S3	0.56*	4.14	**11.68**	8.06	4.7	**19.74**	*24.4*	*0.19*	***0.81***

* Non-significant difference ($p > .05$)

Looking at the vertical displacement in Figure 5.6ace, we notice that for all the subjects the palatalized labial is characterized by raising of Tongue Body. This articulator begins at a relatively low position at VC1 (point 1). It is gradually raised to a substantially higher level at VC2 (point 2) and, particularly, at CV1 (point 3). Further it is lowered again (at CV2, or point 4).

For all the speakers, the highest value of TBy is at the offset of the closure (or its release), point CV1. Notice also that at point CV2 TBy has not yet been lowered to its original position (as at VC1). Unlike /p'/, its plain counterpart, /p/, does not involve a major movement of Tongue Body. The articulator is fairly stable for Subject 1 and Subject 2, while there is some Tongue Body lowering at the CV transition for Subject 3.

For all the subjects the two consonants are significantly different from each other with respect to TBy throughout the examined time period (Table 5.9). Notice, however, that the difference is most substantial at point CV1, the stop release. Combining the differences between /p/ and /p'/ for VC (VC1 and VC2) and CV (CV1 and CV2) transitions allows us to estimate the relative amount of 'information' available about the contrast.[30] Thus, while the VC transition contains from 22 to 40% of the 'information', the CV transition provides from 60 to 78%. Notice that our 'transitions' in this study are only 30 ms long. The difference between the two consonants, however, is likely to be extended much further throughout the following vowel (but not as far into the preceding vowel). The overall, or total, articulatory difference between the initial /p/ and /p'/ in terms of TBy is from 16 to 18 mm.

To summarize, the palatalization of /p'/ is characterized by a Tongue Body raising gesture, which is timed with the stop release. At the same time, the quality of /p/ may optionally involve certain Tongue Body lowering or have no significant movement at all. These articulatory differences between the two stops are highly significant over a substantially long period of time.

Turning to the horizontal displacement (Figure 5.6 bdf), we notice that the palatalized labial is characterized by fronting of Tongue Body. This articulator begins at a relatively back position at VC1. It is gradually brought forward at points VC1 and CV1 (at CV1 for Subject 3), and then retracted again at Point CV2. Speakers vary with respect to the most advanced position of TBx.

[30] Note that this approach involves a certain oversimplification, e.g. assuming that the same amount of information can be equally well recovered from the VC and CV transition, as well as counting the non-significant differences.

It is VC1 for Subject 2, CV1 for Subject 3, and both of these points for Subject 1.[31] Backing of TBx further than the original position at CV2 is due to the quality of the following consonant, /p/ (in *ta papy*).

The articulation of the plain labial /p/ also involves a movement of Tongue Body, but in the opposite direction. While it is relatively neutral at point VC1, it is much further back during the CV transition. Note that the Tongue Body position does not change during the CV transition. This may be also due to the quality of the following plain labial /p/ and high mid vowel [ɨ] (in *papy*).

Again, the differences between the gestures are significant throughout most of the time (except VC1 for Subject 2 and Subject 3) (Table 5.9). The CV transition accounts for 62 to 81% of 'information' about the articulatory differences. The overall articulatory difference between the initial /p/ and /p'/ in terms of TBx is from 23.5 to 25.5 mm, which is higher than for the vertical displacement (not evaluated statistically).

We can summarize this by saying that the palatalization of /p'/ is characterized not only by TB raising, but also by a TB fronting gesture during the stop constriction. At the same time, the plain labial /p/ is characterized by strong Tongue Body retraction, or velarization. These movements of Tongue Body in the opposite directions allow for very significant articulatory differences between the two stops.

To summarize, the results show that in addition to the primary gesture, Lip Aperture, the articulation of /p'/ involves a secondary gesture, or rather two gestures, Tongue Body raising and fronting. I will refer to these two movements together as the 'palatalized gesture'. This gesture is timed with the primary gesture, Lip Aperture, so that the peak of TBy is the beginning of the release of Lip Aperture, and the peak of TBx is during the constriction or at the release of Lip Aperture. The articulation of the plain labial involves the primary, LA, and the secondary,

[31] The difference of TBx during the VC transition of Subject 3 can be explained by the difference in the stimuli. This

Tongue Body backing (or retraction), gesture. Tongue Body can also be lowered during the production of /p/. The secondary gesture is timed at the release of the primary gesture. I will further refer to this movement as the 'velarized gesture'. The maximum difference between plain and palatalized labials is at the CV transition (especially point CV1), i.e. it is timed with the release of the primary gesture.

5.3.2.6.1.2 /t/ vs. /t'/

Figure 5.7 presents the Tongue Body trajectories of /t/ and /t'/ at four points in time for the three speakers. Again, the vertical displacement (TBy) is given in the left column and the horizontal displacement (TBx) is in the right column. For TBy higher values represent higher position of Tongue Body in the oral cavity. For TBx lower values correspond to a more front position of Tongue Body in the oral cavity. The overall Tongue Body differences between /t/ and /t'/ at each point are given in Table 5.10.

Looking at the vertical displacement (Figure 5.7 ace), we notice that for all the subjects, the palatalized coronal is characterized by raising of Tongue Body. This articulator begins at a relatively low position at VC1. It is gradually raised to a substantially higher level at VC1 and CV1, and then lowered again at Point CV2. For all the speakers, there is little difference between TBy values at VC2 and CV1, as well as between the initial and final values.

---

subject was presented with the stimuli 'papy' in the frame *Eto __ opjat'*, where the vowel preceding [t] was often reduced to schwa, while Subject 1 and Subject 2 had an were given 'ta tapy' in the same frame.

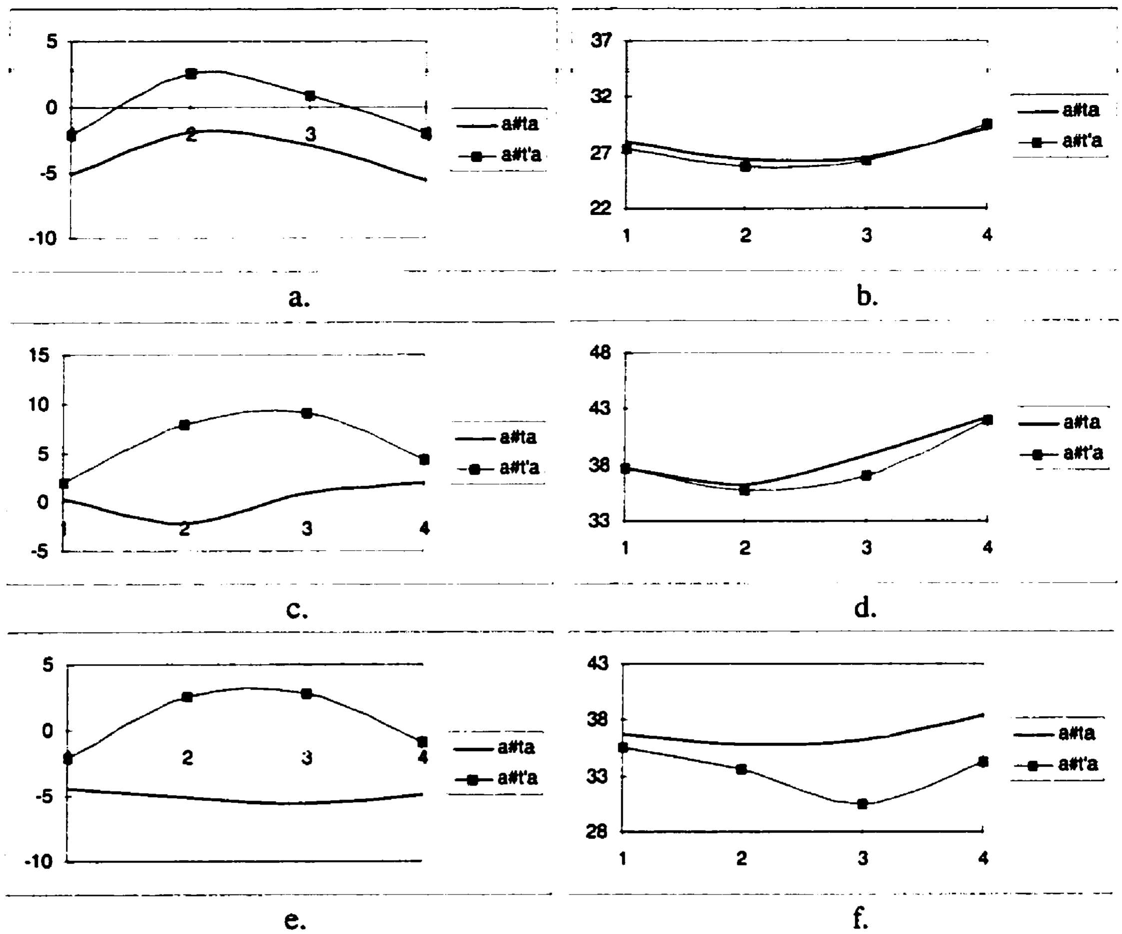


Figure 5.7 Movement of Tongue Body (vertical (ace) and horizontal (bdf)) during the articulation of /t/ and /t'/, based on means at four points in time for Subjects 1 (ab), 2 (cd), and 3 (ef).

Table 5.10 Mean differences between initial /t/ and /t'/

		VC1	VC2	CV1	CV2	Total VC	Total CV	Total	VC, %	CV, %
TBy	S1	3.02	**4.42**	3.84	3.5	**7.44**	7.34	*14.8*	***0.5***	***0.5***
	S2	1.66	**10.14**	8.08	2.36	11.8	**10.44**	*22.2*	***0.53***	*0.47*
	S3	2.3	7.66	**8.36**	4.02	9.96	**12.38**	*22.3*	*0.45*	***0.55***
TBx	S1	0.62[a]	0.68[a]	0.28[a]	0.32[a]	1.3	0.6	*1.9*	***0.68***	*0.32*
	S2	0.08[a]	0.54[a]	1.8[a]	0.28[a]	0.62	2.08	*2.7*	*0.23*	***0.77***
	S3	1.08[a]	2.24[a]	**5.76**	4.22[a]	3.32	**9.98**	*13.3*	*0.25*	***0.75***

[a] Non-significant difference (p > .05)

Speakers vary with respect to Tongue Body movement during plain /t/. Subject 1 raises Tongue Body, but not as high, as for /t'/. The /t/ of Subject 2 involves some Tongue Body lowering at onset of the closure. And /t/ of Subject 3 shows very little difference across the time.

For all the subjects the two consonants are significantly different from each other with respect to TBy throughout the time period examined (Table 5.10). Combining the differences between /t/ and /t'/ for VC (VC1 and VC2) and CV (CV1 and CV2) transitions the amount of information is divided almost equally between VC (from 45 to 53%) and CV (from 47 to 55%) transitions. The total articulatory difference between the initial /t/ and /t'/ in terms of TBy is from 15 to 22 mm.

We can summarize this by saying that the palatalization of /t'/ is characterized by a Tongue Body raising gesture, which occurs at the same time as the primary Tongue Tip gesture. On the other hand, the nature of plain /t/ is not consistently manifested by lowering of Tongue Body. In fact, it may be raised as well.

The statistical results show that labials /p/ and /p'/ are different with respect to TBy throughout the time period examined. Most of the TBy difference between /t/ and /t'/, in contrast, is at both the onset and offset of the constriction, VC2 and CV1 (for Subject 2 and Subject 3). There is little difference between these two points.

Turning to the horizontal displacement (Figure 5.7bdf), both plain and palatalized coronals are characterized by fronting of Tongue Body at VC1 and CV1. Note that the amount of fronting is almost the same for Subject 1 and Subject 2. It is timed with the stop closure. There is significantly more Tongue Body fronting for /t'/ at the CV transition for Subject 3. Thus, the overall articulatory difference between the initial /t/ and /t'/ in terms of TBx is from 2 (not significant) to 13.5 mm (Table 5.10).

The similar behaviour of /t/ and /t'/ in terms of Tongue Body fronting is not surprising, since the primary gesture for both of these consonants is the forward movement of Tongue Tip towards the upper teeth (or alveolar ridge). Tongue Body, being coupled with Tongue Tip, is dragged forward with it regardless of whether the consonant is plain or palatalized.

There is little or no difference between /t/ and /t'/ in horizontal displacement, since the articulation of both consonants involves Tongue Body fronting. This fronting is due to the primary gesture for /t/. For /t'/ fronting is partly due to the primary gesture and partly to the secondary palatalized gesture.

To summarize, palatalized coronal /t'/ is characterized primarily by Tongue Body raising that is timed with the constriction of the primary gesture rather than its release. Tongue Body fronting, also a property of /t/, can only optionally differentiate the two consonants by its degree. The vertical Tongue Body movement during /t/ is less consistent: it may involve certain raising or lowering. Thus, the plain quality of /t/ is not consistently differentiated by Tongue Body. The maximum difference between plain and palatalized coronals tends to be found both at points VC1 and CV2 (Subject 2 and Subject 3; but throughout for Subject 1), or at the onset and offset (release) of the primary constriction.

5.3.2.6.2 Place

We have seen that the articulations of labial and coronal plain and palatalized stops contrast in the movement of Tongue Body in both vertical (raising/lowering) and horizontal (fronting/backing) displacement. The two places of articulation vary in how they employ Tongue Body movement to contrast plain and palatalized consonants. In this section I focus on the place of articulation differences between palatalized (/p'/ vs. /t'/) and plain (/p/ vs. /t/) consonants.

5.3.2.6.2.1 Palatalized

Figure 5.8 presents the Tongue Body trajectories of palatalized stops /p'/ and /t'/ at four points in time for the three speakers. Again, the vertical displacement (TBy) is given in the left column and the horizontal displacement (TBx) in the right column. Recall that for TBy higher values represent higher position of Tongue Body in the oral cavity. For TBx lower values correspond to a more front position of Tongue Body in the oral cavity. The overall Tongue Body differences between /p'/ and /t'/ at each point are given in Table 5.11. Non-significant differences (based on the corresponding main effects of Place and significant Place × Pal interactions) are marked with an asterisk.

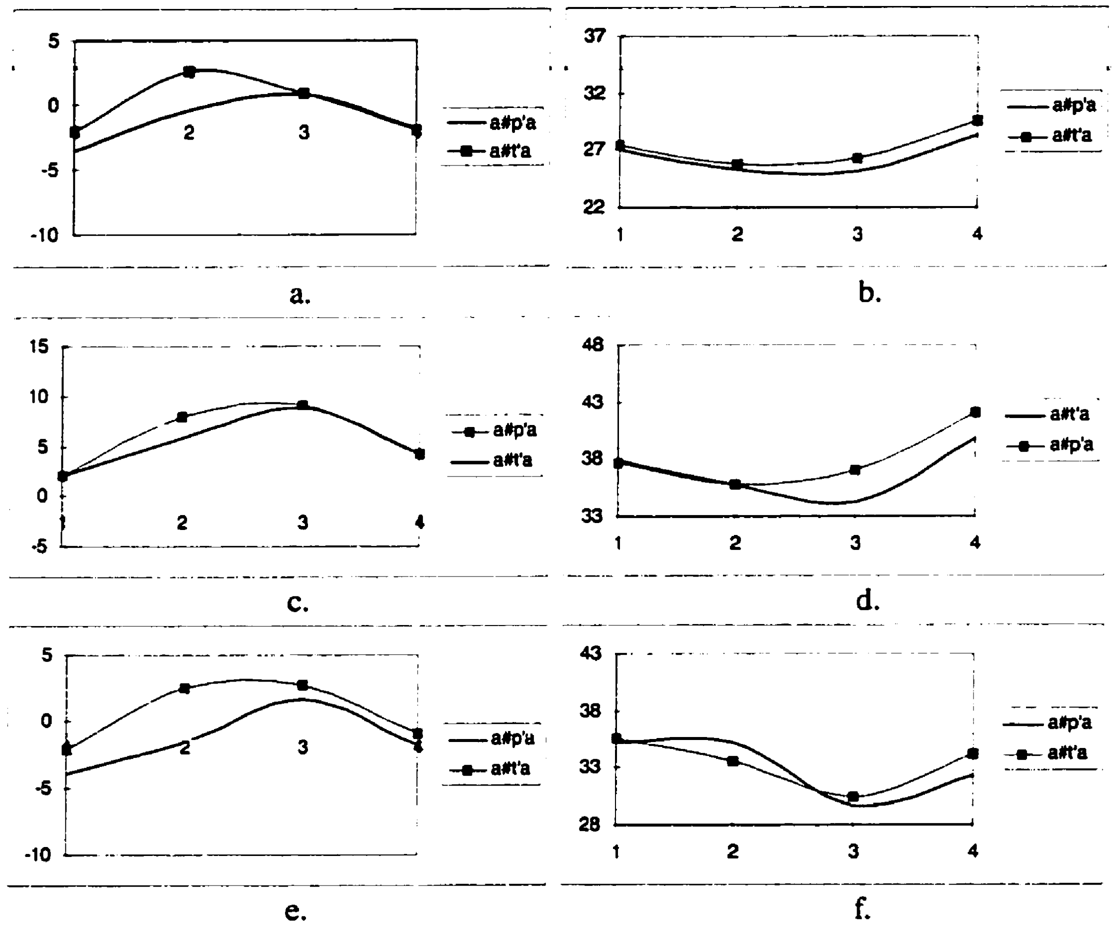


Figure 5.8 Movement of Tongue Body (vertical (ace) and horizontal (bdf)) during the articulation of /p'/ and /t'/, based on means at four points in time for Subjects 1 (ab), 2 (cd), and 3 (ef).

Table 5.11 Mean differences between initial /p'/ and /t'/

		VC1	VC2	CV1	CV2	Total VC	Total CV	Total
TBy	S1	1.8	**2.92**	0.12[a]	0.22[a]	**4.72**	0.34	*5.06*
	S2	0.78[a]	**2.86**	0.16[a]	0.02	**3.64**	0.18	*3.82*
	S3	2.56[a]	**4.1**	1.1	0.18	**6.66**	1.28	*7.94*
TBx	S1	0.4[a]	0.58[a]	1.1[a]	1.24[a]	0.98	**2.34**	*3.32*
	S2	0.02[a]	0.04[a]	**3.66**	3.54[a]	0.06	**7.2**	*7.26*
	S3	0.1[a]	2.64[a]	**3.12**	4.34[a]	2.74	**7.46**	*10.2*

[a] Non-significant difference ($p > .05$)

Looking at the vertical displacement (Figure 5.8ace), while both the palatalized labial and coronal show raising of Tongue Body, the amount of this raising and its timing are not the same for the two places. The Tongue Body of /t'/ is in general raised higher than that of /p'/. While the peak of TBy for /p'/ is clearly at CV1, the peak for /t'/ is during the acoustic closure (or articulatory constriction) (Subject 1 and Subject 2), or at VC2 (Subject 3). Major significant differences between the two consonants are found at Point VC1 (from 3 to 4 ms). /p'/ and /t'/ are not significantly different at CV1 (Subject 1 and Subject 2), or the difference between them is relatively small (1 mm, Subject 3). The total articulatory difference between the initial /p'/ and /t'/ in terms of TBy is from about 4 to 8 mm (a rough estimation, based on both significant and non-significant differences).

In terms of Tongue Body fronting the two places of articulation are similar, with some significant differences only at CV1 for Subject 2 and Subject 3. These two subjects are not consistent in what consonant has more fronting: Subject 2 shows more forward Tongue Body for /t'/ (by about 3.5 mm), while Subject 2 has a more fronted Tongue Body for /p'/ (by about 3 mm). The total articulatory difference between the initial /p'/ and /t'/ in terms of TBx is from about 3.5 (not significant) to 10 mm.

In sum, /p'/ and /t'/ are similar since their articulation involves Tongue Body raising and fronting. However, the corresponding Tongue Body gestures are not identical: the Tongue Body of /t'/ is higher than that of /p'/. The palatalized gesture of /p'/ is timed later than the same gesture for /t'/. The difference between the two consonants in terms of Tongue Body raising is greater during the VC transition. There are also some less consistent differences in terms of Tongue Body fronting, mainly during the CV transition.

5.3.2.6.2 Plain

Figure 5.9 presents the Tongue Body trajectories of plain stops /p/ and /t/ at four points in time for the three speakers. Again, the vertical displacement (TBy) is given in the left column and the horizontal displacement (TBx) is in the right column. For TBy higher values represent a higher position of Tongue Body in the oral cavity. For TBx lower values correspond to a more front position of Tongue Body in the oral cavity. The overall Tongue Body differences between /p/ and /t/ at each point are given in Table 5.12.

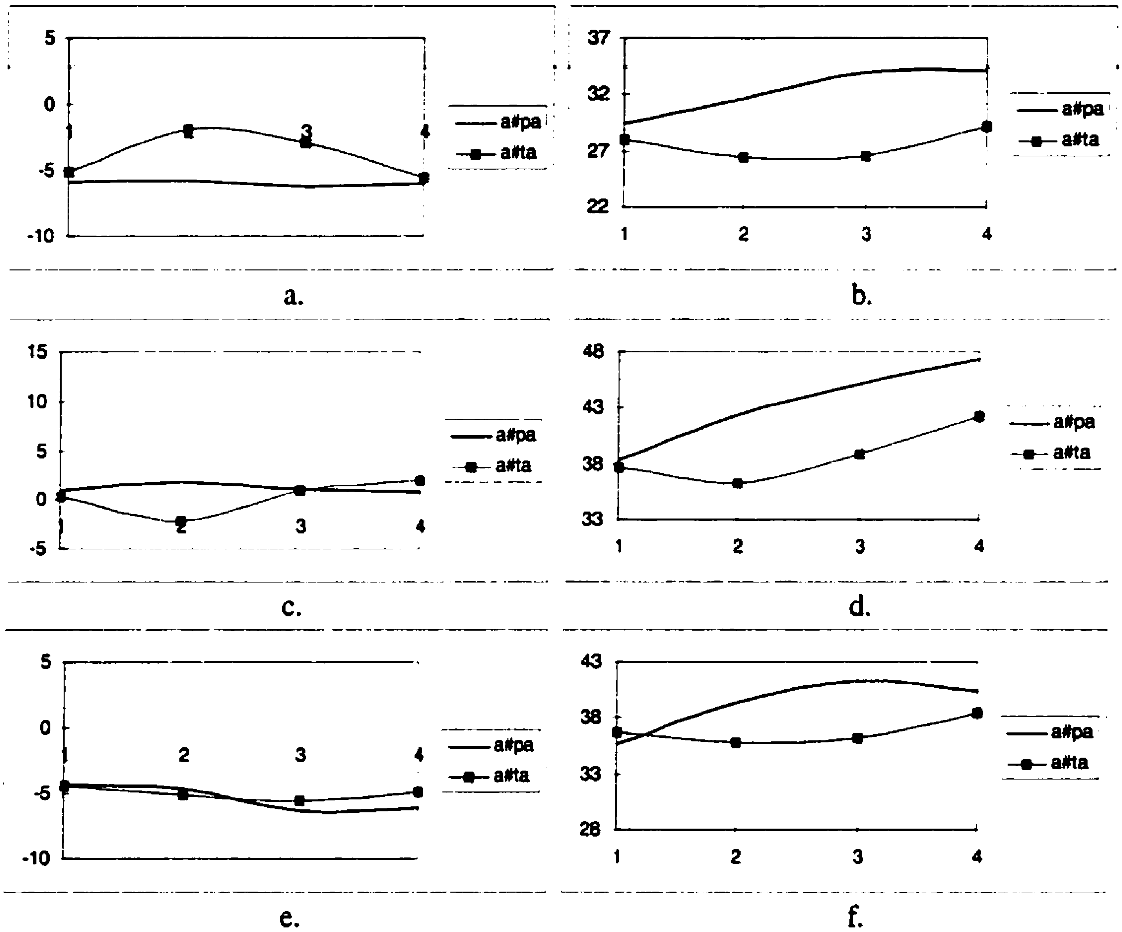


Figure 5.9 Movement of Tongue Body (vertical (ace) and horizontal (bdf)) during the articulation of /p/ and /t/, based on means at four points in time for Subjects 1 (ab), 2 (cd), and 3 (ef).

Table 5.12 Mean differences between initial /t/ and /t'/

		VC1	VC2	CV1	CV2	Total VC	Total CV	Total
TBy	S1	0.74*	**3.94**	3.3	0.52*	**4.68**	3.82	*8.5*
	S2	0.56*	**4.06**	0.1*	1.12	**4.62**	1.22	*5.84*
	S3	0.06*	0.46*	0.76	**1.18**	0.52	**1.94**	*2.46*
TBx	S1	1.4	5.18	**7.42**	4.96	6.58	**12.38**	*19*
	S2	0.66*	6.1	**6.26**	5.1	6.76	**11.36**	*18.1*
	S3	1.52*	3.52	**5.14**	1.96	5.04	**7.1**	*12.1*

* Non-significant difference (p > .05)

We can see from Figure 5.8ace that speakers vary in their TBy place differences: while TBy for /t/ is higher than for /p/ for Subject 1 (at VC2 and CV1) and Subject 2 (at CV1 and CV2), the Tongue Body for /t/ is lower than for /p/ for Subject 2 (at VC2). These articulatory differences are not very large, from about 2.5 to 8.5 mm (a rough estimation, based on both significant and non-significant differences).

The difference between the two consonants is much more substantial in terms of the horizontal displacement, TBx. For all the speakers, /p/ involves Tongue Body backing, or retraction, while /t/ shows consistent Tongue Body fronting. The difference between the two is more apparent during the CV transition, since the Tongue Body retraction gesture of /p/ is timed with the stop release. The total articulatory difference between the initial /p/ and /t/ in terms of TBx is from about 12 to 19 mm.

To summarize, in addition to their primary gestures, /p/ and /t/ are different from each other in terms of the direction of the horizontal (and optionally, vertical) movement of Tongue Body. While /p/ involves backing of Tongue Body (velarization), /t/ shows fronting of the same articulator.

We have seen that the palatalized quality of a consonant varies depending on the place of articulation. In other words, there are important articulatory asymmetries between /p'/ and /t'/ with respect to their contrast with plain consonants. First, while /p'/ is different from /p/ both in terms of Tongue Body raising/lowering (TBy) and Tongue Body fronting/backing (TBx) (especially in TBx), /t'/ is differentiated from /t/ only (or largely) with respect to Tongue Body raising/lowering (TBy). Second, while the maximum differences between /p'/ and /p/ are at the CV transition (especially Point CV1), those between /t'/ and /t/ are distributed more equally between VC and CV transitions. Third, the relative magnitude of the palatalized labial is lower than that of the palatalized coronal. There are also differences between plain consonants /p/ and /t/. While the articulation of /p/ is characterized by Tongue Body backing, that of /t/ involves Tongue Body fronting. Neither /p/ nor /t/ is consistent across speakers with respect to Tongue Body raising/lowering.

Why do we find these differences between the palatalized gestures of /p'/ and /t'/? It is very likely that this is due to the nature of the articulators involved in the production of two consonants (Louis Goldstein, p.c.). Lips and Tongue Body, involved in the articulation of /p'/, are relatively independent of each other and the timing of the two gestures is not restricted articulatorily. The production of /t'/, on the other hand, involves two articulators, Tongue Tip (TT) and Tongue Body, that are two parts of the same major articulator, the tongue. In other words, they are articulatorily coupled (Kelso et al. 1986). The movement of one of them is likely to affect the trajectory of the other. There are two conflicting targets during production of a palatalized /t'/: to make a constriction (closed) with Tongue Tip at the teeth (fronting and some raising) and to make another constriction (narrow) with Tongue Body against the hard palate (raising and fronting). These two movements are, to a large degree, in the same direction. During the production of /t/ the movement of Tongue Tip on its own raises and fronts Tongue Body to a certain extent. The addition of the palatalized gesture is likely to lead to a higher (and possibly more front) position of Tongue Body. This explains the difference between /t'/ and /p'/ in the degree of Tongue Body raising. The movement of Tongue Tip also affects the timing of Tongue Body. As a result of the fronting and raising of Tongue Tip, Tongue Body is raised and fronted to its peak earlier, i.e. during the closure rather than at the release of Tongue Tip (as /p'/).

The asymmetries between the plain /p/ and /t/ in terms of the velarized gesture can also be partly attributed to the articulatory differences between the articulators. While Lips do not preclude Tongue Body from maximum retraction during the production of /p/, the fronting movement of Tongue Tip drags Tongue Body forward and makes it harder to perform a velarized gesture.[32]

To conclude, the plain-palatalized contrast is strongly affected by the articulatory differences of gestures involved in the production of both palatalized (/p'/ and /t'/) and plain (/p/ and /t/) stops.

[32] Instead, TB can be somewhat lowered (Subject 2), but this may be due to individual differences in articulatory apparatus.

5.3.2.6.3 Environment

Here I address the Tongue Body differences between the same consonants in two different environments, initial and final. I focus on the palatalized stops /p'/ and /t'/, since their positional differences are of particular importance.

5.3.2.6.3.1 /p'/

Figure 5.10 presents the Tongue Body trajectories of the palatalized stop /p'/ in initial and final environments at four points in time for the three speakers. The overall Tongue Body differences between /#p'/ and /p'#/ at each point are given in Table 5.13. Non-significant differences are based on the corresponding main effects of Env and significant Place × Env and Pal × Env interactions.

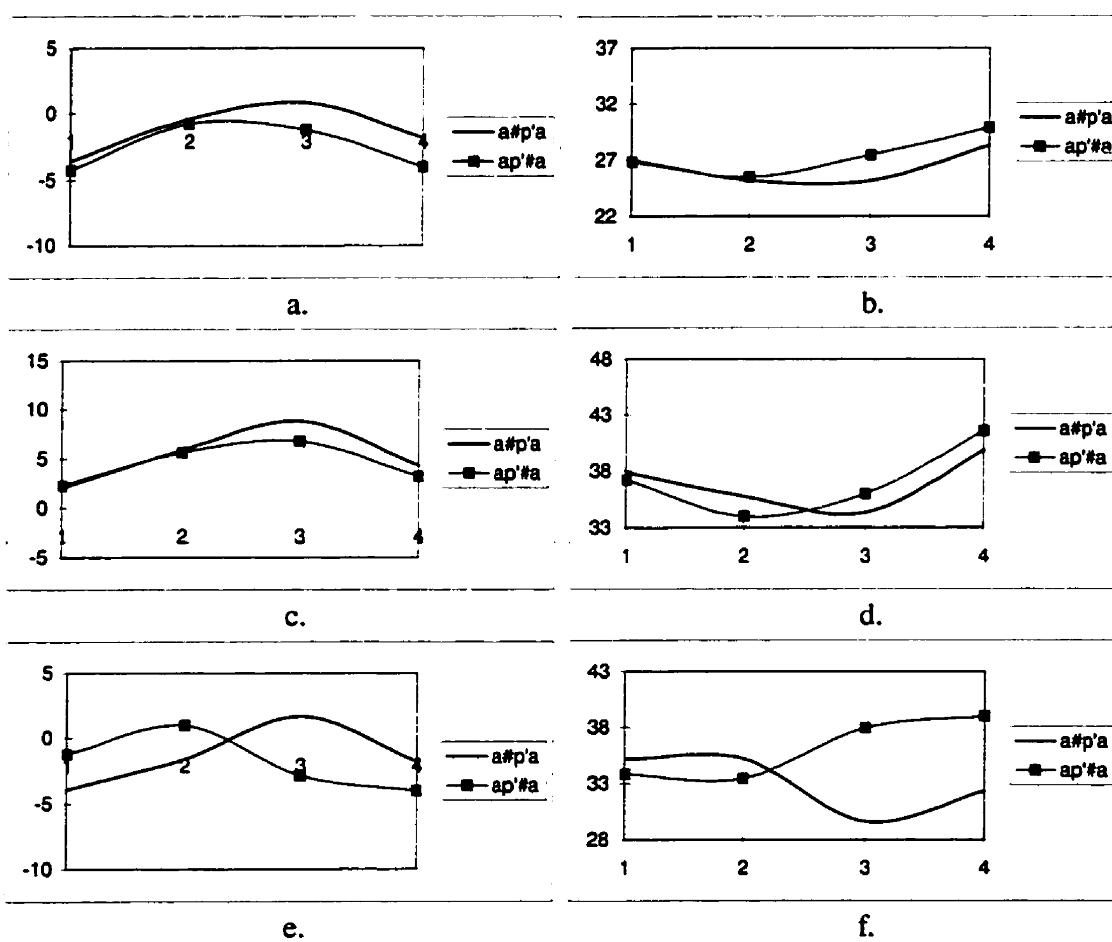


Figure 5.10 Movement of Tongue Body (vertical (ace) and horizontal (bdf)) during the articulation of initial and final /p'/, based on means at four points in time for Subjects 1 (ab), 2 (cd), and 3 (ef).

Table 5.13 Mean differences between initial and final /p'/

		VC1	VC2	CV1	CV2	Total VC	Total CV	Total
TBy	S1	0.7•	0.38•	**2.08•**	2.18•	1.08	**4.26**	*5.34*
	S2	0.26•	0.26•	**2.04•**	1.06•	0.52	**3.1**	*3.62*
	S3	2.72•	2.6	**4.5**	2.22•	5.32	**6.72**	*12*
TBx	S1	0.16•	0.26•	**2.3**	1.58	0.42	**3.88**	*4.3*
	S2	0.64•	**1.74**	1.68•	**1.74•**	2.38	**3.42**	*5.8*
	S3	1.32•	1.76•	**8.3**	6.64	3.08	**14.94**	*18*

• Non-significant difference ($p > .05$)

Looking at the vertical displacement (Figure 5.10ace), the mean values for the two environments are not the same: the TBy gesture in final position is somewhat lower than in initial position. These differences, while consistent across the speakers, however, are not significant for Subject 1 and Subject 2 (however, see the discussion of other interactions below). Only for Subject 3 the difference is significant at VC2. Notice that for this subject, the peak of the final TBy is timed with the beginning of the closure rather than with its release, as for the initial /p'/. In other words, the Tongue Body gesture in final position for Subject 3 is reduced in magnitude and shifted forward in time. The results for the other subjects can be interpreted only as a tendency in this direction. The total articulatory difference between the initial and final /p'/ in terms of TBy is from about 3 to 6.5 mm (a rough estimation, based on both significant and non-significant differences).

There are more substantial differences in terms of horizontal displacement (Figure 5.13bdf). The final Tongue Body gesture is either reduced in magnitude (Subject 1 and Subject 3), or shifted forward (all the subjects). Again, Subject 3 shows a major difference between the two positions (at VC2, CV1, and CV2). The total articulatory difference between the initial and final /p'/ in terms of TBx is from about 4.5 up to 18 mm.

To summarize, the differences between palatalized labials in two environments are not as large as those between the two different consonants overall. The final /p'/ shows a tendency for reduction and shift of its Tongue Body gesture (and particularly, TBx), compared with the initial /p'/.

5.3.2.6.3.2 /t'/

Figure 5.11 presents the Tongue Body trajectories of the palatalized stop /t'/ in initial and final environments at four points in time for the three speakers. The overall Tongue Body differences between initial and final /t'/ at each point are given in Table 5.14.

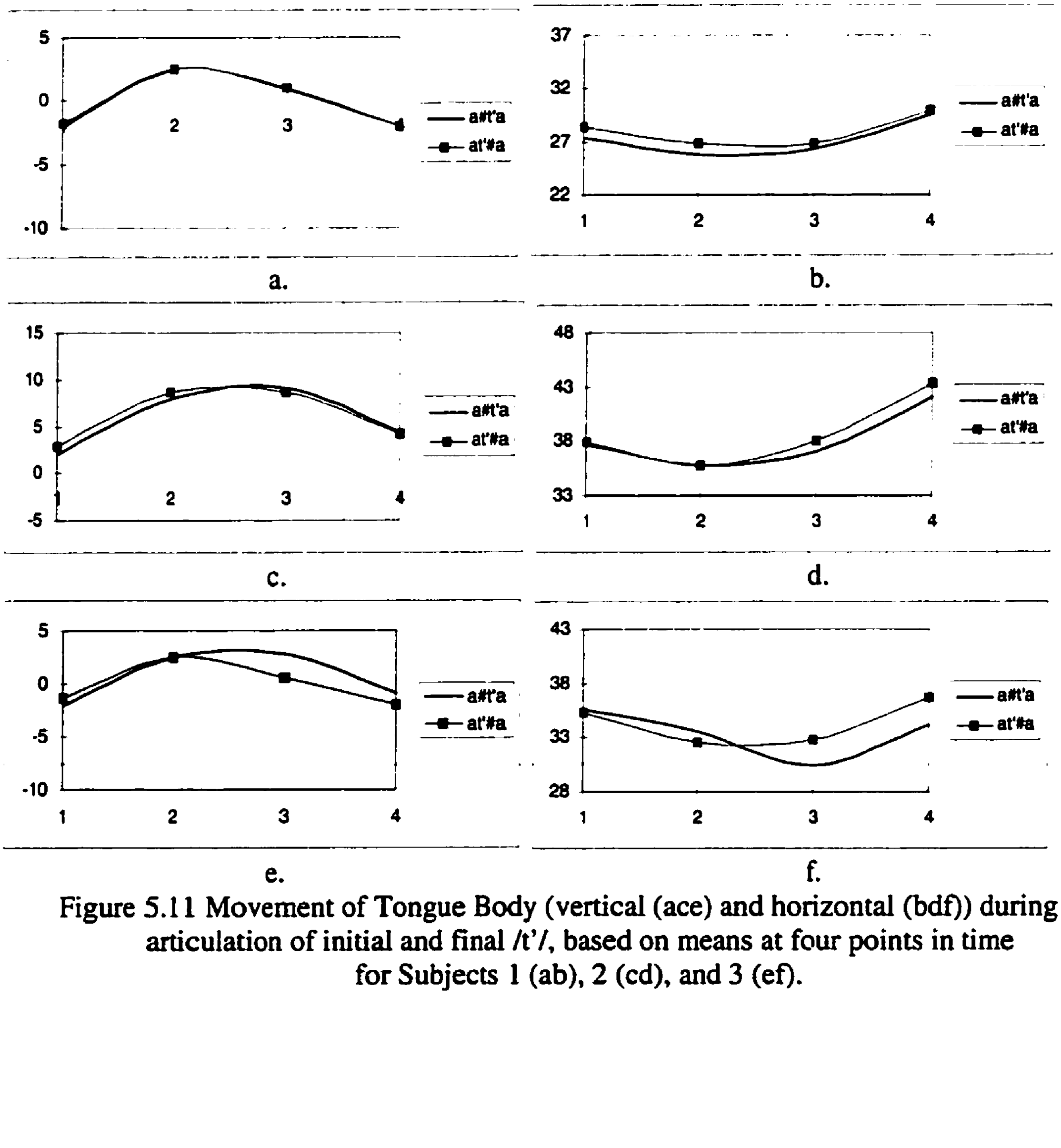


Figure 5.11 Movement of Tongue Body (vertical (ace) and horizontal (bdf)) during the articulation of initial and final /t'/, based on means at four points in time for Subjects 1 (ab), 2 (cd), and 3 (ef).

Table 5.14 Mean differences between initial and final /t'/

		VC1	VC2	CV1	CV2	Total VC	Total CV	Total
TBy	S1	0.3*	0.04*	0.06*	0.06*	0.34	0.12	*0.46*
	S2	0.8*	0.8*	0.44*	0.06*	1.6	0.5	*2.1*
	S3	0.72*	0.04*	**5.64**	1.06*	0.76	**6.7**	*7.46*
TBx	S1	0.98*	1*	0.54	0.5	1.98	1.04	*3.02*
	S2	0.16*	0.04*	0.9*	1.36*	0.2	**2.26**	*2.46*
	S3	0.34*	1.02	**2.34**	2.46	1.36	**4.8**	*6.16*

*Non-significant difference ($p > .05$)

The Tongue Body configuration of the palatalized coronal stop does not differ very much from environment to environment either in terms of Tongue Body raising or with respect to Tongue Body fronting (Subject 1 and Subject 2) (however, see the discussion of other interactions below). Only Subject 3 shows some reduction and forward shift of the Tongue Body peak. The total articulatory difference between the initial and final /t'/ in terms of TBy is from about .5 to 7.5 mm (mostly non-significant). The difference in terms of TBx is from about 2.5 to 6 mm (partly non-significant).

In sum, the palatalized coronal is less affected by position than the palatalized labial. This is even more true for the plain consonants, /p/ and /t/, that do not show any positional reduction.

5.3.2.6.4 Environment, place and palatalization

A number of interactions also point to asymmetries in terms of environment, place, and palatalization (see sections 5.3.2.2-5.3.2.4). Thus, labials are more sensitive to the Environment (#labial vs. labial#) than coronals (the significant Place × Env interaction: TBx at VC2 for Subject 2, TBx at CV1 and CV2 for Subject 3). Environment particularly affects palatalized /p'/ (the Place × Pal × Env interaction, TBx at CV1 for Subject 3).

Palatalized consonants show higher values in the initial environment and lower values in final position, while the plain consonants remain relatively stable (the Pal × Env interaction: TBy at CV1 for Subject 2, TBx at CV1 and CV2 for Subject 3).

There is a difference in terms of Palatalization between initial consonants but not between final segments (the Pal × Env interaction: TBx at CV2 for Subject 2, TBy at CV2, TBx at CV1 and CV2 for Subject 3) and between final labial and coronal segments (the Place × Pal × Env interaction: TBx at CV1 for Subject 3).

In sum, Environment is an important factor, although less prominent than Palatalization and Place. The palatalized labial /p'/ is primarily affected by Environment: its Tongue Body gesture (and, particularly, TBx) is somewhat reduced and/or shifted forward.

How can we explain the effect of Environment? A study of English glide gestures in different syllable positions (Gick 1999) shows that reduction of glide gestures is very common in the coda environment. The reduction of the palatalized gesture in the coda position can be compared to the palatal glide /j/ reduction in codas. The relative independence of Tongue Body from Lips in /p'/ leads to a natural decrease in the magnitude of the secondary gesture. It does not affect the Tongue Body of /t'/ to the same extent, since it is coupled with the primary gesture Tongue Tip. More reduction and shift of Tongue Body found in the results for Subject 3 can be also attributed to the overall faster speech rate (see section 5.5.2.2).

5.3.2.7 Summary

The results of the experiment revealed that the articulator that plays a crucial role in distinguishing plain and palatalized consonants is the Tongue Body. We have seen that consonants in question vary in the magnitude of the Tongue Body, in the direction of movement of this articulator, as well as in its timing.

5.3.2.8 Conclusion

In this section I investigated articulatory differences between four stops in terms of Tongue Body movement. The results show that this is the major articulatory correlate of the plain-palatalized contrast. The production of palatalized consonants always involves Tongue Body raising and fronting, or the 'palatalized gesture'. This gesture, however, is not identical for /p'/

and /t'/. The plain consonants were found to have some lowering and backing of Tongue Body, the 'velarized gesture', however, it is not very consistent (especially for /t/). The palatalized gesture (particularly, of /p'/) was found to be susceptible to environment.

In the next section we examine the articulatory properties of the same consonants in clusters (coda) and compare them to the single segments.

5.3.3 Consonants in clusters

This section presents the results for Tongue Body difference between plain and palatalized consonants in the preconsonantal coda position.

5.3.3.1 Overall results

In this section I give a general overview of the results. The results by major groups of stimuli are discussed in more detail in later sections.

Tables 5.15-5.17 present means and standard deviations for Tongue Body magnitude during the VC transition by dimension (vertical, TBy and horizontal, TBx) and by subject (based on the average of 5 tokens per subject). The consonants of interest (C_1) and their contexts (C_2) are given in bold.

In the case of the utterance *tap papy* as pronounced by Subject 1 (Table 5.15a), at time VC1 the Tongue Body had the magnitude of -4.08 (.95) mm (vertical displacement: left columns). The first number here stands for the mean value and the number in parentheses is the standard deviation. Later, at point VC2, the Tongue Body was lowered to -4.92 (1.33) mm (by less than 1 mm). At the same time on the horizontal displacement (Table 5.15b), Tongue Body is backed from 29.06 (1.89) mm at VC1 to 33.10 at VC2 (by 4 mm).

Table 5.15 Tongue Body magnitude at two points: Subject 1

a. Vertical displacement, TBy

	VC1		VC2				VC1		VC2	
	mean	SD	mean	SD			mean	SD	mean	SD
tap papy	-4.08	0.95	-4.92	1.33		tat papy	-5.82	0.93	-2.84	0.96
tap tapy	-3.44	1.55	-4.50	1.04		tat tapy	-4.98	0.95	-2.18	0.84
tap p'apy	-3.76	1.67	-3.34	0.83		tat p'apy	-4.58	1.62	-2.68	1.85
tap t'apy	-4.04	0.97	-3.64	0.88		tat t'apy	-4.28	1.76	-1.24	2.45
tap' papy	-3.96	0.51	0.08	0.86		tat' papy	-1.38	2.95	1.54	0.76
tap' tapy	-2.92	2.06	-0.44	1.70		tat' tapy	-2.12	1.51	1.56	1.84
tap' p'apy	-3.48	0.91	-0.52	1.10		tat' p'apy	-1.64	0.66	1.58	0.63
tap' t'apy	-3.48	0.98	-2.08	0.80		tat' t'apy	-2.13	1.54	1.98	1.26

b. Horizontal displacement, TBx

	VC1		VC2				VC1		VC2	
	mean	SD	mean	SD			mean	SD	mean	SD
tap papy	29.06	1.89	33.10	0.64		tat papy	29.22	0.82	27.30	1.15
tap tapy	28.46	0.65	30.88	0.69		tat tapy	28.16	1.13	26.88	0.95
tap p'apy	29.90	1.27	32.72	1.71		tat p'apy	27.92	0.77	26.48	0.61
tap t'apy	29.54	0.71	32.28	0.48		tat t'apy	27.96	1.15	26.64	1.32
tap' papy	27.36	1.00	25.86	0.61		tat' papy	28.94	1.03	26.80	0.80
tap' tapy	27.22	0.61	27.02	0.62		tat' tapy	27.82	0.65	26.44	0.72
tap' p'apy	27.62	0.50	26.34	0.50		tat' p'apy	27.90	0.93	26.14	0.80
tap' t'apy	28.46	1.11	29.40	2.21		tat' t'apy	28.43	0.54	26.80	0.62

Table 5.16 Tongue Body magnitude at two points: Subject 2

a. Vertical displacement, TBy

	VC1		VC2				VC1		VC2	
	mean	SD	mean	SD			mean	SD	mean	SD
tap papy	1.96	0.82	2.66	0.48		tat papy	-1.76	1.78	-3.40	1.04
tap tapy	1.35	0.99	0.15	1.49		tat tapy	-0.44	1.53	-2.38	1.03
tap p'apy	1.82	0.6	3.18	1.03		tat p'apy	0.28	1.4	-1.06	2.52
tap t'apy	1.32	1.49	2.78	1.33		tat t'apy	0.78	1.94	3.22	4.68
tap' papy	1.87	0.12	5.40	0.35		tat' papy	2.96	0.9	7.60	1.52
tap' tapy	1.96	0.78	5.04	0.74		tat' tapy	1.66	1.76	5.26	2.16
tap' p'apy	1.68	0.73	5.42	0.39		tat' p'apy	2.94	0.13	8.36	0.72
tap' t'apy	1.62	0.88	5.28	1.19		tat' t'apy	2.72	0.88	7.92	0.81

b. Horizontal displacement, TBx

	VC1		VC2				VC1		VC2	
	mean	SD	mean	SD			mean	SD	mean	SD
tap papy	38.50	1.4	41.92	1.32		tat papy	37.80	2.21	36.58	1.63
tap tapy	39.15	0.42	40.73	1.61		tat tapy	37.88	1.18	36.40	1.49
tap p'apy	39.18	1.4	40.28	1.72		tat p'apy	38.88	1.27	37.96	2.11
tap t'apy	39.08	0.69	40.72	0.52		tat t'apy	37.98	1.03	35.24	1.66
tap' papy	38.30	0.78	35.90	1.06		tat' papy	38.16	2.57	35.48	2
tap' tapy	38.02	0.86	36.60	1.29		tat' tapy	37.48	1.62	34.22	1.04
tap' p'apy	38.04	0.68	35.38	0.77		tat' p'apy	39.36	1.49	36.26	0.58
tap' t'apy	39.04	1.26	37.14	2.08		tat' t'apy	38.66	0.42	36.46	0.62

Table 5.17 Tongue Body magnitude at two points: Subject 3

a. Vertical displacement, Tby

	VC1		VC2				VC1		VC2	
	mean	SD	mean	SD			mean	SD	mean	SD
tap papy	-3.64	.83	-3.40	.76		tat papy	-3.70	.53	-3.26	.68
tap tapy	-4.36	.39	-5.06	.19		tat tapy	-4.30	.14	-3.18	.59
tap p'apy	-2.70	.29	-0.28	.05		tat p'apy	-3.58	.72	-1.80	.35
tap t'apy	-3.10	.87	-0.48	.03		tat t'apy	-2.38	.82	0.90	.97
tap' papy	-2.08	.12	0.00	.77		tat' papy	-0.58	.95	3.58	.2
tap' tapy	-1.84	.75	0.32	.51		tat' tapy	-1.30	.4	1.93	.01
tap' p'apy	0.30	.06	2.77	.66		tat' p'apy	-0.22	.82	4.36	.19
tap' t'apy	-1.94	.01	0.62	.93		tat' t'apy	-0.58	.65	3.46	.86

b. Horizontal displacement, TBx

	VC1		VC2				VC1		VC2	
	mean	SD	mean	SD			mean	SD	mean	SD
tap papy	36.06	.93	38.50	.94		tat papy	34.42	.32	34.00	.05
tap tapy	36.50	.19	36.64	.34		tat tapy	35.40	.27	33.32	.73
tap p'apy	34.45	.04	33.80	.72		tat p'apy	34.90	.36	32.45	.37
tap t'apy	35.04	.88	35.42	.65		tat t'apy	35.08	.91	32.66	.56
tap' papy	35.25	.49	34.85	.96		tat' papy	34.52	.98	31.56	.32
tap' tapy	34.40	.78	33.14	.51		tat' tapy	34.33	.64	31.73	.26
tap' p'apy	33.90	.11	32.13	.77		tat' p'apy	35.00	.82	31.82	.48
tap' t'apy	34.40	.77	33.06	.23		tat' t'apy	35.86	.8	32.76	.5

The following sections (5.3.3.2-5.3.3.4) present statistical results by subjects and their summary (section 5.3.3.5). These are necessary to determine what factors are significant. Please turn to section 5.3.3.6 for the linguistically relevant discussion of the facts.

5.3.3.2 Results (statistics): Subject 1

Table 5.18 shows the ANOVA results for Tongue Body magnitude (Tby, left and TBx, right) for the first (coda) consonant in a cluster, as pronounced by Subject 1. The significant results ($p < .05$) are given in bold. The main effects and significant interactions are further described point by point.

Table 5.18 Tongue Body magnitude: Subject 1 (statistics)

	TBy				TBx			
	VC1		VC2		VC1		VC2	
	F	*p*	*F*	*p*	*F*	*p*	*F*	*p*
Place	.73	.395	**55.15**	**.000**	.53	.470	**177.57**	**.000**
Pal	**28.29**	**.000**	**159.47**	**.000**	**13.61**	**.000**	**140.82**	**.000**
EnvPlace	.25	.616	.06	.810	1.14	.289	.781	.380
EnvPal	.25	.616	.58	.448	.72	.401	1.94	.169
Place × Pal	**17.51**	**.000**		n.s.	**12.18**	**.001**	**112.93**	**.000**
Place × EnvPlace		n.s.	3.79	.056		n.s.		n.s.
Place × EnvPal		n.s.		n.s.	**9.31**	**.003**	**8.38**	**.005**
Pal × EnvPlace		n.s.		n.s.		n.s.	**16.90**	**.000**
Pal × EnvPal		n.s.	**5.36**	**.024**		n.s.		n.s.
EnvPlace × EnvPal		n.s.		n.s.	**5.10**	**.027**	**8.51**	**.005**
Place × Pal × EnvPlace		n.s.		n.s.		n.s.	**12.19**	**.001**

5.3.3.2.1 Tongue Body: vertical displacement (TBy)

In this section I examine the statistical results for the vertical displacement of the Tongue Body at the different points of measurement taken point by point. Note that these and other results are discussed in section 5.3.3.6.

Point VC1

Pal was significant [$F(1,80) = 28.29$, $p = .000$], while Place, EnvPlace, and EnvPal were not significant.

There was a significant Place × Pal interaction [$F(1,80) = 17.51$, $p = .000$]. This interaction was explored with the Tukey HSD test. The reason for the interaction is that the factor Pal was significant for level two (coronal, /t/ vs. /t'/) of the factor Place ($p = .001$) but not for level one (labial, /p/ vs. /p'/) of this factor ($p = .853$). Also, the factor Place was significant for level two (palatalized, /p'/ vs. /t'/) of the factor Pal ($p = .004$) but not for level one (plain. /p/ vs. /t/) of this factor ($p = .097$).

Point VC2

Pal [F(1,80) = 159.47, p = .000] and Place [F(1,80) = 55.15, p = .000] were significant. EnvPlace and EnvPal were not significant.

There was a significant Pal × EnvPal interaction [F(1,80) = 5.36, p = .024]. The post-hoc test revealed that the factor Pal was significant for both level one (plain, CC vs. C'C) (p = .000) and level two (palatalized, CC' vs. C'C') (p = .000) of the factor EnvPal. The factor EnvPal was significant neither for level one (plain, CC vs. CC') (p = .141) nor level two (palatalized, C'C vs. C'C') (p = .693) of the factor Pal. The interaction can be explained by the asymmetry of the mean values (Means: CC -3.61, CC' -2.73, C'C .69, C'C' .24).

In sum, Pal was significant throughout the VC transition (at points VC1 and VC2). Place was significant at Point VC2. EnvPlace and EnvPal were not significant.

5.3.3.2.2 Tongue Body: horizontal displacement (TBx)

In this section I examine the statistical results for the horizontal displacement of the Tongue Body at the different points of measurement taken point by point.

Point VC1

Pal was significant [F(1,80) = 13.61, p = .000]. Place, EnvPlace and EnvPal were not significant.

There was a significant Place × Pal interaction [F(1,80) = 12.18, p = .001]. The reason for the interaction is that the factor Pal was significant for level one (labial, /p/ vs. /p'/) of the factor Place (p = .000) but not for level two (coronal, /t/ vs. /t'/) of this factor (p = .999). Also, the factor Place was significant for level one (plain, /p/ vs. /t/) of the factor Pal (p = .021) but not for level two (palatalized, /p'/ vs. /t'/) of this factor (p = .218).

There was a significant Place × EnvPal interaction [F(1,80) = 9.31, p = .003]. The reason for the interaction is that the factor Place was significant for level two (palatalized, labial-C' vs. coronal-C') of this factor (p = .046), but not for level one (plain, labial-C vs. coronal-C) of the factor EnvPal (p = .362). Also, the factor EnvPal was significant for level one (labial, labial-C vs. labial-C') of the factor Place (p = .037) but not for level two (coronal, coronal-C vs. coronal-C') of this factor (p = .407). (Means: labial-C 28.03, labial-C' 28.88, coronal-C 28.54, coronal-C' 28.05).

There was a significant EnvPlace × EnvPal interaction [F(1,80) = 5.10, p = .027]. The factor EnvPal was significant neither for level two (coronal, _t vs. _t') of the factor EnvPlace (p = .137) nor level one (labial, _p vs. _p') of this factor (p = .750). Also, the factor EnvPlace was well above the significance level for both level one (plain, _p vs. _t) (p = .097) and level two (palatalized, _p' vs. _t') (p = .836) of the factor EnvPal. asymmetry (Means: _p 28.65, _p' 28.34, _t 27.92, _t' 28.60).

Point VC2

Pal [F(1,80) = 140.82, p = .000] and Place were significant [F(1,80) = 177.57, p = .000]. EnvPlace and EnvPal were not significant.

There was a significant Place × EnvPal interaction [F(1,80) = 8.38, p = .005]. The reason for the interaction is that the factor EnvPal was significant for level one (labial, labial-C vs. labial-C') of the factor Place (p = .018) but not for level two (coronal-C vs. coronal-C') of this factor (p = .713). The factor Place was significant for both level one (plain, labial-C vs. coronal-C) of the factor EnvPal (p = .000) and level two (palatalized, labial-C' vs. coronal-C') of this factor (p = .000).

There was a significant Pal × EnvPlace interaction [F(1,80) = 16.90, p = .000]. The reason for the interaction is that the factor EnvPlace was significant for level two (palatalized, C'-labial vs. C'-coronal) of the factor Pal (p = .004), but not for level one (plain, C-labial vs. C-coronal) of this factor (p = .113). The factor Pal was significant for both level one (labial, C-labial vs. C'-labial) (p = .000) and level two (coronal, C-coronal vs. C'-coronal) (p = .000) of the factor EnvPlace.

There was a significant EnvPlace × EnvPal interaction [F(1,80) = 8.51, p = .005]. The factor EnvPal was significant for level two (coronal, _t vs. _t') of the factor EnvPlace (p = .017) but not for level one (labial, _p vs. _p') of this factor (p = .704). Also, the factor EnvPlace was significant for level two (palatalized, _p' vs. _t') of the factor EnvPal (p = .044) but not for level one (plain, _p vs. _t) of this factor (p = .481). (Means: _p 28.27, _p' 27.92, _t 27.81, _t' 28.78).

There was a significant Place × Pal interaction [F(1,80) = 112.93, p = .000]. The reason for the interaction is that the factor Pal was significant for level one (labial, /p/ vs. /p'/) of the factor Place (p = .000) but not for level two (coronal, /t/ vs. /t'/) of this factor (p = .818). Also, the factor Place was significant for level one (plain /p/ vs. /t/) of the factor Pal (p = .000) but not for level two (palatalized, /p'/ vs. /t'/) of this factor (p = .236).

There was a significant Place × Pal × EnvPlace interaction [F(1,80) = 12.19, p = .001]. The reason for the interaction is that the factor Place was significant for some combinations of levels of factors EnvPlace and Pal (plain-labial (p = .000), plain-coronal (p = .000), palatalized-coronal (p = .018)) but not for others (palatalized-labial (p = .992)). Also, the factor Pal was significant for some combinations of levels of factors Place and EnvPlace (labial-labial (p = .000), labial-coronal (p = .000)) but not for others (coronal-labial (p = .982), coronal-coronal (p = 1.000)). EnvPlace was significant for some combinations of levels of factors Pal and EnvPlace (/p'/ (p = .001)) but not for others (/p/ (p = .080), /t/ (p = 1.000), /t'/, (p = 1.000)).

In sum, Pal was significant throughout the VC transition. Place was significant at Point VC2. EnvPlace and EnvPal were not significant.

5.3.3.3 Results (statistics): Subject 2

Table 5.19 shows the ANOVA results for Tongue Body magnitude (TBy and TBx separately) for the first (coda) consonant in a cluster, as pronounced by Subject 2. The significant results (p < .05) are given in bold. Marginally significant (.05 < p < .06) results are italicized.

Table 5.19 Tongue Body magnitude: Subject 2 (statistics)

	TBy				TBx			
	VC1		VC2		VC1		VC2	
	F	*p*	*F*	*p*	*F*	*p*	*F*	*p*
Place	**4.50**	**.038**	2.12	.150	1.716	.195	**63.01**	**.000**
Pal	**33.42**	**.000**	**224.03**	**.000**	.343	.560	**78.43**	**.000**
EnvPlace	.14	.712	.09	.769	.153	.697	.796	.376
EnvPal	2.97	.090	**24.00**	**.000**	**4.31**	**.042**	.408	.525
Place × Pal	**26.36**	**.000**	**45.73**	**.000**		n.s.	**34.59**	**.000**
Place × EnvPlace		n.s.	*3.87*	*.054*		n.s.	**5.02**	**.028**
Place × EnvPal	**5.70**	**.020**	**6.95**	**.010**		n.s.	*3.70*	*.059*
Pal x EnvPal		n.s.	**6.04**	**.017**		n.s.		n.s.
Pal x EnvPlace		n.s.		n.s.		n.s.	**4.00**	**.050**
EnvPal x EnvPlace		n.s.	**6.17**	**.016**		n.s.		n.s.
Place × Pal × EnvPlace	**4.59**	**.036**	**12.10**	**.001**		n.s.		n.s.

5.3.3.3.1 Tongue Body: vertical displacement (TBy)

In this section I examine the statistical results for the vertical displacement of the Tongue Body at the different points of measurement taken point by point.

Point VC1

Pal was significant [F(1,80) = 33.42, p = .000]. So was Place [F(1,80) = 4.50, p = .038]. EnvPal and EnvPlace were not significant.

There was a significant Place x EnvPal interaction [F(1,80) = 4.59, p = .020]. The reason for the interaction is that the factor Place was significant for level one (plain, labial-C vs. coronal-C) of the factor Pal (p = .012) but not for level two (palatalized, labial-C' vs. coronal-C') of this factor (p = .998). The factor EnvPal was significant for level two (coronal, coronal-C vs. coronal-C') of the factor Place (p = .025) but not for level one (labial, labial-C vs. labial-C') of this factor (p = .965).

There was a significant Place x Pal interaction [F(1,80) = 26.36, p = .000]. The reason for the interaction is that the factor Pal was significant for level two (coronal, /t/ vs. /t'/) of the factor Place (p = .000) but not for level one (labial, /p/ vs. /p'/) of this factor (p = .968). Also, the factor Place was significant for level one (plain /p/ vs. /t/) of the factor Pal (p = .000) but not for level two (palatalized, /p'/ vs. /t'/) of this factor (p = .155).

There was a significant Place × Pal × EnvPlace interaction [F(1,80) = 4.59, p = .036]. The reason for the interaction is that the factor Place was significant for some combinations of levels of factors EnvPlace and Pal (plain-labial (p = .000)) but not for others (palatalized-labial (p = .338), plain-coronal (p = .350), palatalized-coronal (p = .994)). The factor Pal was significant for some combinations of levels of factors Place and EnvPlace (coronal-labial (p = .000), coronal-coronal (p = .006)) but not for others (labial-labial (p = .990), labial-coronal (p = .988)). EnvPlace was not significant for all combinations of levels of factors Pal and EnvPlace (/p/ (p = .963), /p'/ (p = 1.000), /t/ (p = .662), /t'/, (p = .829)).

Point VC2

Pal [F(1,80) = 224.03, p = .000] and EnvPal [F(1,80) = 24.00, p = .000] were significant. Place and EnvPlace were not significant.

There was a significant Place × Pal interaction [F(1,80) = 45.73, p = 000]. The reason for the interaction is that the factor Pal was significant for level one (labial, /p/ vs. /p'/) of the factor Place (p = .000) but not for level two (coronal, /t/ vs. /t'/) of this factor (p = .000). The factor Place was significant for both level one (plain /p/ vs. /t/) of the factor Pal (p = .000) and level two (palatalized, /p'/ vs. /t'/) of this factor (p = .000). The interaction can be explained by the asymmetry of the mean values (Means: /p/ 2.19, /p'/ 5.29, /t/ -.91, /t'/ 7.29).

There was a significant Place × EnvPal interaction [F(1,80) = 6.04, p = .017]. The reason for the interaction is that the factor EnvPal was significant for level two (coronal, coronal-C vs. coronal-C') of the factor Place (p = .000) but not for level one (labial, labial-C vs. labial-C') (p = .386) of this factor. Also, the factor Place was significant for level one (plain, labial-C vs. coronal-C) of the factor EnvPal (p = .026) but not for level two (palatalized, labial-C' vs. coronal-C') of this factor (p = .838).

There was a significant Pal × EnvPal interaction [F(1,80) = 6.04, p = .017]. The reason for the interaction is that the factor EnvPal was significant for level one (plain, CC vs. CC') of the factor Place (p = .000) but not for level two (palatalized, C'C vs. C'C') of this factor (p = .319).

The factor Pal was significant for both level one (plain, CC vs. C'C) of the factor EnvPal (p = .000) and level two (palatalized, CC' vs. C'C') of this factor (p = .000).

There was a significant EnvPlace × EnvPal interaction [F(1,80) = 6.17, p = .016]. The reason for the interaction is that the factor EnvPal was significant for level two (coronal, _t vs. _t') of this factor (p = .000) but not for level one (labial, _p vs. _p') of the factor EnvPlace (p = .329). The factor EnvPlace was significant neither for level one (plain, _p vs. _t) of the factor EnvPal (p = .212) nor level two (palatalized, _p' vs. _t') of this factor (p = .416).

There was a significant Place × Pal × EnvPlace interaction [F(1,80) = 12.0, p = .001]. The reason for the interaction is that the factor Place was significant for some combinations of levels of factors EnvPlace and Pal (plain-labial (p = .000), palatalized-labial (p = .024)) but not for others (plain-coronal (p = .860), palatalized-coronal (p = .558)). The factor Pal was significant for all combinations of levels of factors Place and EnvPlace (labial-labial (p = .032), labial-coronal (p = .000), coronal-labial (p = .000), coronal-coronal (p = .001)). EnvPlace was significant for some combinations of levels of factors Pal and EnvPlace (/t/ (p = .018)) but not for others (/p/ (p = .536), /p'/ (p = 1.000), /t'/ (p = .593)).

In sum, Pal was significant throughout the VC transition. Place was significant at point VC1 and EnvPal at point VC2. EnvPlace was not significant.

5.3.3.3.2 Tongue Body: horizontal displacement (TBx)

In this section I examine the statistical results for the horizontal displacement of the Tongue Body at the different points of measurement taken point by point for Subject 2.

Point VC1

EnvPal was significant [F(1,80) = 4.31, p = .042]. Pal, Place, and EnvPlace were not significant and was not significant. There were no significant interactions.

Point VC2

Pal and Place were significant (Pal: [F(1,80) = 78.43, p = .000], Place: [F(1,80) = 63.01, p = .000]). EnvPal and EnvPlace were not significant.

There was a significant Place × Pal interaction [F(1,80) = 35.59, p = .000]. The reason for the interaction is that the factor Pal was significant for level one (labial, /p/ vs. /p'/) of the factor Place (p = .000) but not for level two (coronal, /t/ vs. /t'/) of this factor (p = .163). Also, the factor Place was significant for level one (plain /p/ vs. /t/) of the factor Pal (p = .000) but not for level two (palatalized, /p'/ vs. /t'/) of this factor (p = .471).

There was a significant Place × EnvPlace interaction [F(1,80) = 5.02, p = .028]. The factor EnvPlace was significant neither for level one (labial, labial-labial vs. labial-coronal) of the factor Place (p = .777) nor level two (coronal, coronal-labial vs. coronal-coronal) of this factor (p = .130). At the same time, the factor Place was significant for both level one (labial, labial-labial vs. coronal-labial) of the factor EnvPlace (p = .001) and level two (coronal, labial-coronal vs. coronal-coronal) of this factor (p = .000). The interaction can be explained by the asymmetry of the mean values (Means: labial-labial 38.37, labial-coronal 38.80, coronal-labial 36.57, coronal-coronal 35.58).

There was a significant Pal × EnvPlace interaction [F(1,80) = 4.00, p = .050]. The factor Pal was significant for both level one (labial, C-labial vs. C'-labial) of the factor EnvPlace (p = .000) and level two (coronal, C-coronal vs. C'-coronal) of this factor (p = .000). At the same time, the factor EnvPlace was significant neither for level one (plain C-labial vs. C-coronal) of the factor Pal (p = .182) nor level two (palatalized, C'-labial vs. C'-coronal) of this factor (p = .862). The interaction can be explained by the asymmetry of the mean values (Means: C-labial 39.19, C-coronal 38.27, C'-labial 35.76, C'-coronal 36.11).

In sum, Pal and Place were significant at Point VC2. EnvPal was significant at Point VC1. EnvPlace was not significant.

5.3.3.4 Results (statistics): Subject 3

Table 5.20 shows the ANOVA results for Tongue Body magnitude (TBy and TBx separately) for the first (coda) consonant in a cluster, as pronounced by Subject 3.

Table 5.20 Tongue Body magnitude: Subject 3 (statistics)

	TBy				TBx			
	VC1		VC2		VC1		VC2	
	F	*p*	*F*	*p*	*F*	*p*	*F*	*p*
Place	1.16	.287	**15.04**	**.000**	.033	.856	**45.99**	**.000**
Pal	**59.72**	**.000**	**128.28**	**.000**	2.45	.122	**38.30**	**.000**
EnvPlace	2.04	.158	1.36	.248	.88	.351	.022	.882
EnvPal	**9.05**	**.004**	**39.40**	**.000**	.711	.402	**14.38**	**.000**
Place × Pal		n.s.	**6.83**	**.011**		n.s.	**6.78**	**.011**
Place × EnvPal		n.s.		n.s.	**6.07**	**.016**	**9.40**	**.003**
Pal × EnvPal		n.s.	**7.04**	**.010**		n.s.	**6.78**	**.011**
EnvPlace × EnvPal		n.s.		n.s.		n.s.	**9.35**	**.003**
Pal × EnvPlace × EnvPal		n.s.	3.82	.055		n.s.		n.s.

5.3.3.4.1 Tongue Body: vertical displacement (TBy)

In this section I examine the statistical results for the vertical displacement of the Tongue Body at the different points of measurement taken point by point for Subject 3.

Point VC1

Pal [F(1,80) = 59.72, p = .000] and EnvPal [F(1,80) = 9.05, p = .004] were significant. Place and EnvPlace were not significant. There were no significant interactions.

Point VC2

Pal was significant [F(1,80) = 128.28, p = .000]). So were Place [F(1,80) = 15.04, p = .000]) and EnvPal [F(1,80) = 39.40, p = .000]). EnvPlace was not significant.

There was a significant Place × Pal interaction [F(1,80) = 6.78, p = .011]. The reason for the interaction is that the factor Place was significant for level two (palatalized, /p'/ vs. /t'/) of the factor Pal (p = .000) but not for level one (plain, /p/ vs. /t/) of this factor (p = .807). The factor Pal was significant for both level one (labial, /p/ vs. /p'/) of the factor Place (p = .000) and level two (coronal, /t/ vs. /t'/) of this factor (p = .000).

There was a significant Pal × EnvPal interaction [F(1,80) = 7.04, p = .010]. The reason for the interaction is that the factor EnvPal was significant for level one (plain, C-C vs. C-C') of the factor Pal (p = .000) but not for level two (palatalized, C'-C vs. C'-C') of this factor (p = .60). The factor Pal was significant for both level one (plain, C-C vs. C'-C) of the factor EnvPal (p = .000) and level two (palatalized, C-C' vs. C'-C') of this factor (p = .000).

In sum, Pal was significant throughout the VC transition. Place was significant at point VC2. EnvPal was significant at both VC1 and VC2. EnvPlace was not significant.

5.3.3.4.2 Tongue Body: horizontal displacement (TBx)

In this section I examine the statistical results for the horizontal displacement of the Tongue Body at the different points of measurement taken point by point.

Point VC1

None of the main effects were significant. At Point VC1 there was a significant Place × EnvPal interaction [F(1,80) = 6.07, p = .016]. The factor EnvPal was significant neither for level one (labial, labial-C vs. labial-C') of the factor Place (p = .100) nor level two (coronal, coronal-C vs. coronal-C') of this factor (p = .658). The factor Place was significant neither for level one (plain, labial-C vs. coronal-C) of the factor EnvPal (p = .245) nor level two (palatalized, labial-C' vs. coronal-C') of this factor (p = .381). The reason for the interaction is the asymmetric relations of the means (Means: labial-C 35.55, labial-C' 34.45, coronal-C 34.67, coronal-C' 35.21).

Point VC2

Pal and Place were significant (Pal: [F(1,80) = 38.30, p = .000], Place: [F(1,80) = 45.99, p = .000]). EnvPal was also significant [F(1,80) = 14.38, p = .000]). EnvPlace was not significant.

There was a significant Place × Pal interaction [F(1,80) = 6.78, p = .011]. The reason for the interaction is that the factor Pal was significant for level one (labial, /p/ vs. /p'/) of the factor Place (p = .000) but not for level two (coronal, /t/ vs. /t'/) of this factor (p = .062). The factor Place was significant for both level one (plain, /p/ vs. /t/) of Pal factor (p = .022) and level two (palatalized, /p'/ vs. /t'/) of this factor (p = .000).

There was a significant EnvPlace × EnvPal interaction [F(1,80) = 9.35, p = .003]. The factor EnvPal was significant for level one (labial, _p vs. _p') of the factor EnvPlace (p = .000) but not for level two (coronal, _t vs. _t') of this factor (p = .956). The factor EnvPlace was significant neither for level one (plain, _p vs. _t) of the factor EnvPal (p = .113) nor level two (palatalized, _p' vs. _t') of this factor (p = .178).

There was a significant Place × EnvPal interaction [F(1,80) = 9.40, p = .003]. The reason for the interaction is that the factor EnvPal was significant for level one (labial, labial-C vs. labial-C') of the factor Place (p = .000) but not for level two (coronal, coronal-C vs. coronal-C') of this factor (p = .959). The factor Place was not significant for level one (plain, labial-C vs. coronal-C) of the factor EnvPal (p = .160) and marginal for level two (palatalized, labial-C' vs. coronal-C') of this factor (p = .053).

There was a significant Pal × EnvPal interaction [F(1,80) = 6.78, p = .011]. The factor EnvPal was significant for level one (plain, CC vs. CC') of the factor Place (p = .000) but not for level two (coronal, C'C vs. C'C') of this factor (p = .833). The factor Pal was significant for level one (plain, CC vs. C'C) of the factor EnvPal (p = .000) but not for level two (palatalized, CC' vs. C'C') of this factor (p = .062).

In sum, Pal and Place were significant at VC2. EnvPal was also significant at Point VC2. EnvPlace was not significant.

5.3.3.5 Summary

In this section I summarize the statistical results for Tongue Body magnitude.

Tongue Body: vertical displacement (TBy)

Pal is significant throughout the VC transition for all three subjects. Place is more variable: it is significant either at the first point, VC1 (Subject 2), or at the second point, VC2 (Subject 1 and Subject 3). EnvPal is either significant at both VC1 and VC2 (Subject 3), at VC2 only

(Subject 2), or not significant altogether (Subject 1). EnvPlace is not significant for any of the speakers.

Among the interactions, Place × Pal is significant throughout the transition (Subject 2) only at VC1 (Subject 1) or only at VC2 (Subject 3). The Pal × EnvPal interaction is significant at VC2 for all the speakers. The Place × EnvPal interaction is significant at both VC1 and VC2 (Subject 2), only at VC2 (Subject 3) or is not significant (Subject 1). Other significant interactions, EnvPal × EnvPlace (VC2), and Place × Pal × EnvPlace (VC1 and VC2) are found only for Subject 2.

Tongue Body: horizontal displacement (TBx)

Pal is either significant throughout the VC transition (Subject 1) or only at VC2 (Subject 2 and Subject 3). Place is significant only at VC2 (all the subjects). EnvPal is significant either at VC1 (Subject 2), VC2 (Subject 3) or is not significant (Subject 1). EnvPlace is not significant for any of the speakers.

Among the interactions, Place × Pal is significant throughout the transition (Subject 1), or only at VC2 (Subject 2 and Subject 3). The Place × EnvPal interaction is significant at both VC1 and VC2 (Subject 1 and Subject 3), or only at VC2 (Subject 2). The EnvPlace × Pal interaction is significant at VC2 (all the speakers).

The EnvPal × EnvPlace interaction is significant at VC1 and VC2 (Subject 1) at VC2 (Subject 3), or not significant (Subject 2). The interaction Place × Pal × EnvPlace (VC1 and VC2) is significant only for Subject 1, and Pal × EnvPal holds only for Subject 3.

Overall, Pal is highly significant for all the speakers, most often throughout the transition. It is followed by Place, and further, by EnvPal. EnvPlace is never significant.

5.3.3.6 Discussion

In this section I discuss the articulatory (Tongue Body) differences between consonants and environments that are of phonetic interest for our study. I further compare the coda consonants in clusters to the single segments.

5.3.3.6.1 Palatalization

Recall that unlike single consonants, preconsonantal stops do not have the CV transition. The main acoustic source of information about the Tongue Body trajectories is the VC transition (in addition to burst, if available). Here I address the Tongue Body differences between the preconsonantal plain and palatalized labials (/p/ vs. /p'/) and coronals (/t/ vs. /t'/) as a whole (averaged by environment), leaving the discussion of environment for section 5.3.3.6.3.

5.3.3.6.1.1 /p/ vs. /p'/

Recall that in section 5.3.2.6.1.1 we determined that, with singleton consonants, palatalized labials are characterized by raising and fronting of Tongue Body, while plain labials show backing of this articulator. The differences between the plain and palatalized gestures are significant throughout the period examined, VC and CV transitions (30 ms each), but are at a maximum during the CV transition in singletons. Now, in clusters, the question can again be asked: Are these two consonant types well differentiated before consonants?

Figure 5.12 presents the Tongue Body trajectories of /p/ and /p'/ for the three speakers, averaged by environments (i.e. _p, _p', _t, and _t'). The trajectories are based on measurements at two points in time: Point 1 (VC1), at 30 ms before the acoustic closure of the stop, and Point 2 (VC2), or at the onset of the closure. The vertical displacement (TBy) is given in the left column and the horizontal displacement (TBx) in the right column. For TBy higher values represent higher position of Tongue Body in the oral cavity. For TBx lower values correspond to a more front position of Tongue Body in the oral cavity.

The overall Tongue Body differences between /p/ and /p'/ at each point are given in Table 5.21. Non-significant differences (based on the corresponding main effects of Pal and significant Place × Pal interactions) are marked with an asterisk. Note that the differences between VC1 and VC2 were not evaluated statistically.

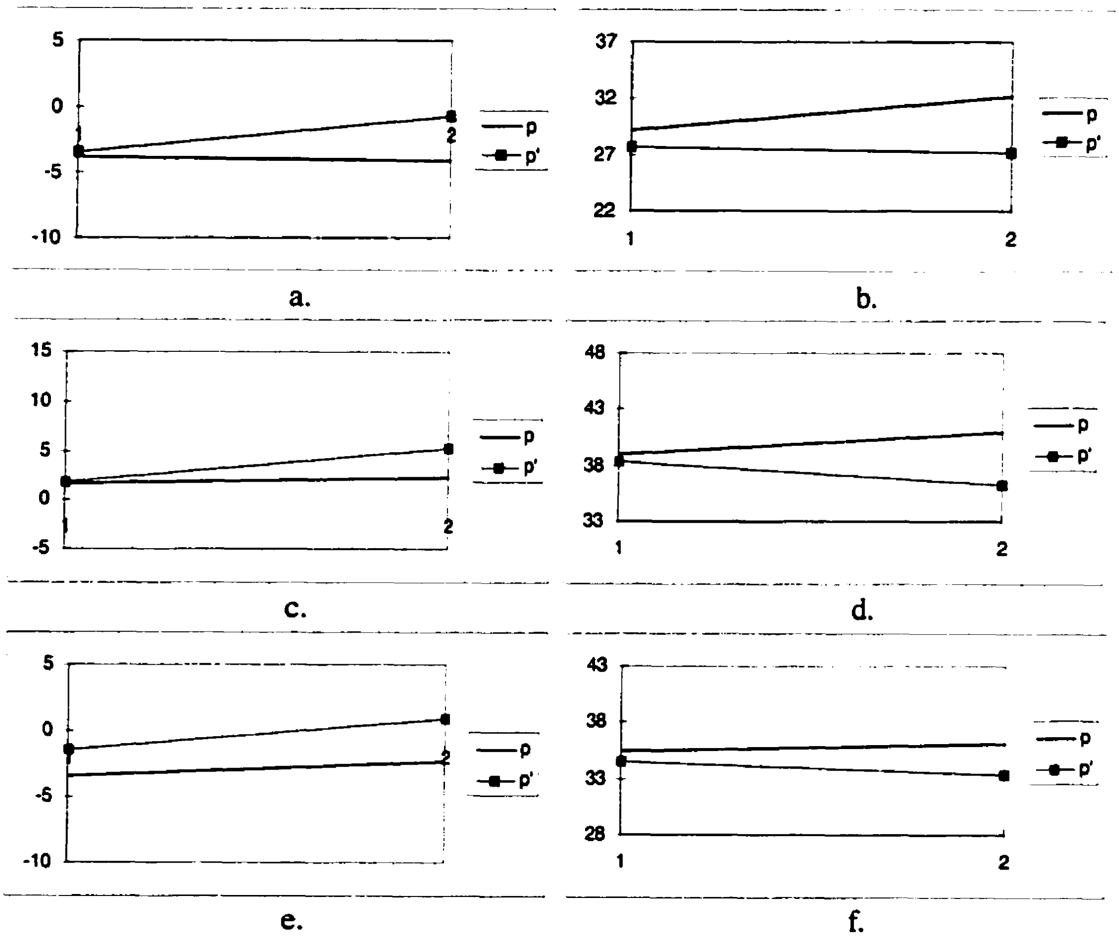


Figure 5.12 Movement of Tongue Body (vertical (ace) and horizontal (bdf)) during the articulation /p/ and /p'/ as first consonant in a cluster, based on means at two points in time for Subjects 1 (ab), 2 (cd), and 3 (ed).

Table 5.21 Mean differences between /p/ and /p'/

		VC1	VC2	Total
TBy	S1	0.51*	3.77	*4.28*
	S2	0.37*	3.03	*3.40*
	S3	2.09	3.16	*5.24*
TBx	S1	1.89	5.69	*7.57*
	S2	0.88*	4.79	*5.67*
	S3	1.03*	2.77	*3.80*

* Not significant difference ($p < .05$)

Looking at the vertical displacement in Figure 5.11ace, while there is no difference (or little difference for Subject 3) between the two movements at the beginning of the transition VC1 (unlike with single consonants), the two gestures are significantly apart at the onset of the closure, VC2. /p'/, as we saw with single consonants, is characterized by Tongue Body raising. The plain /p/ does not involve a major movement of Tongue Body. It should be noted, however, that the amount of this raising is less than we saw during the VC transition of the final single consonant.

Recall that Subject 3 showed a significant temporal shift of the palatalized gesture in the final position. The same happens here: for this speaker the difference between /p/ and /p'/ is significant throughout the transition. The overall, or total, articulatory difference (both at VC1 and VC2) between the preconsonantal /p/ and /p'/ in terms of TBy is from about 3.5 to 5 mm (cf. single /p/ and /p'/ in the final position: 4.5 to 6.5 mm). Note that this is an averaged value for all environments. This can be affected by the quality of the following consonant, as discussed in section 5.3.3.6.3.

Turning to the horizontal displacement (Figure 5.11bdf), we notice that the Tongue Body trajectories for /p'/ and /p/ go in opposite directions. As before, we observe fronting of Tongue Body for /p'/ and backing of this articulator for /p/. For all the speakers the differences are significant at VC2, and for Subject 1 they hold throughout the transition. The overall articulatory difference between /p/ and /p'/ in terms of TBx is from 4 to 7.5 mm (cf. single /p/ and /p'/ in final position: 7.5 to 10 mm).

To summarize, the plain and palatalized labials before consonants are articulatorily distinct from each other (however, not always throughout the transition). This is manifested in both raising/fronting during /p'/ and backing during /p/.

5.3.3.6.1.2 /t/ vs. /t'/

As we learned in section 5.3.2.6.1.2, the two coronal consonants as singletons were distinguished primarily by Tongue Body raising during /t'/ and its relative absence during /t/. Both stops involved fronting of Tongue Body. The differences between the two Tongue Body trajectories were spread almost equally throughout the examined period, VC and CV transitions. Below we determine the degree of articulatory contrast between the two consonants in the preconsonantal position.

Figure 5.13 presents the Tongue Body trajectories of /t/ and /t'/ during the VC transition, averaged by environment. Again, the vertical displacement (TBy) is given in the left column and the horizontal displacement (TBx) is in the right column. For TBy higher values represent higher position of Tongue Body in the oral cavity. For TBx lower values correspond to a more front position of Tongue Body in the oral cavity. The overall Tongue Body differences between /t/ and /t'/ at each point are summarized in Table 5.22.

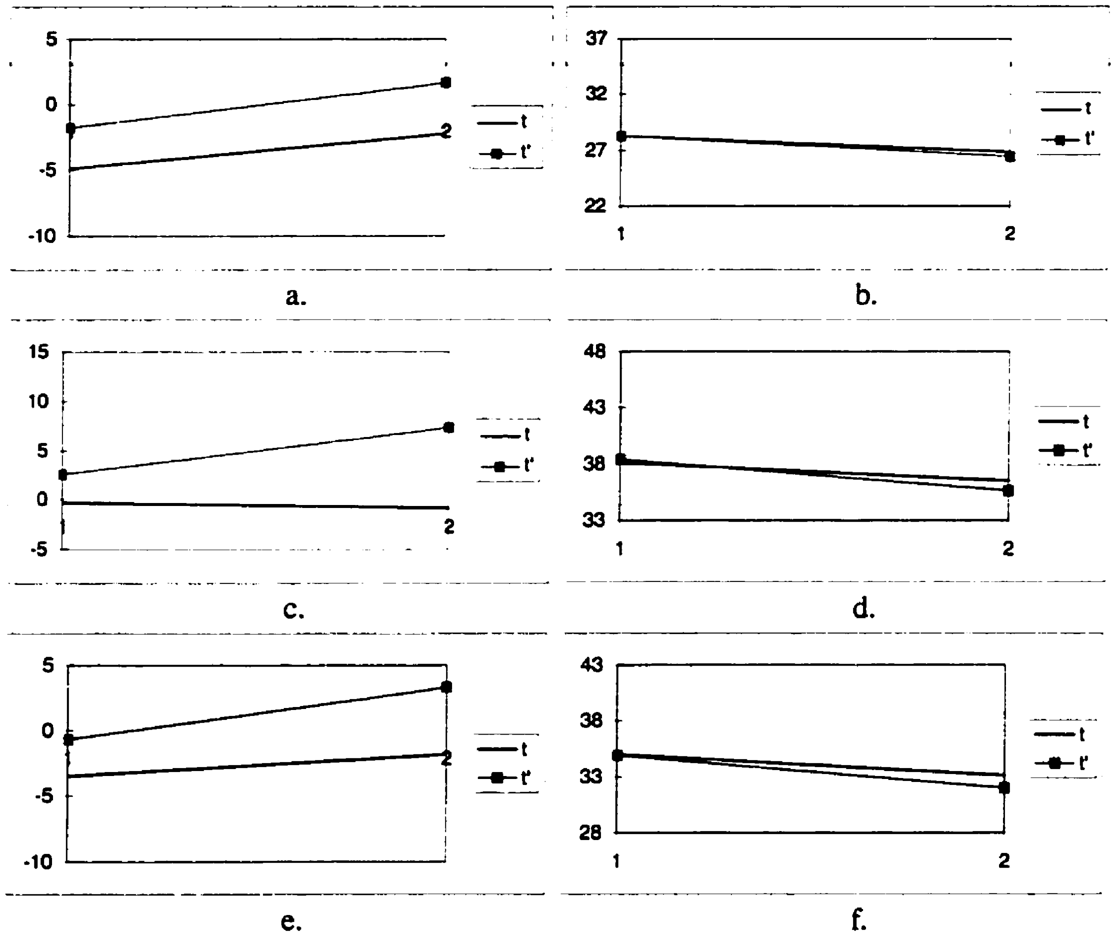


Figure 5.13 Movement of Tongue Body (vertical (ace) and horizontal (bdf)) during the articulation /t/ and /t'/ as first consonant in a cluster, based on means at two points in time for Subjects 1 (ab), 2 (cd), and 3 (ed).

Table 5.22 Mean differences between /t/ and /t'/

		VC1	VC2	Total
TBy	S1	3.10	3.90	*7.00*
	S2	2.86	8.19	*11.05*
	S3	2.82	5.17	*7.99*
TBx	S1	0.28*	0.36*	*0.64*
	S2	0.48*	1.55*	*2.03*
	S3	0.51*	1.19*	*1.70*

* Not significant difference ($p < .05$)

At the vertical displacement in Figure 5.13ace, the two trajectories are significantly different at both points for all three speakers. The palatalized /t'/ shows substantial Tongue Body raising. As before, speakers vary in the production of /t/, however, not exactly the same as with single consonants. Recall that the /t/ of Subject 1 (section 5.3.2.6.1.2) showed Tongue Body raising (the slope of 3.24 mm in initial position); it was almost neutral for Subject 3 (-.74 mm), and Subject 2 had a substantial Tongue Body lowering (-2.58 mm). This is not the case for the preconsonantal /t/ as pronounced by Subject 2 (-.62 mm) and Subject 3 (1.66 mm). As we will see later (section 5.3.3.6.3), the raising by Subject 3 and lack of lowering by Subject 2 are due to the strong effect of a following palatalized segment. The overall articulatory difference between the preconsonantal /t/ and /t'/ in terms of TBy is from about 7 to 11 mm (cf. single /t/ and /t'/ in the final position: 7 to 14 mm).

On the horizontal displacement (Figure 5.13bdf) the trajectories are statistically the same. This is also consistent with the findings for single consonants. The overall difference between /t/ and /t'/ in terms of TBx is very minor, from .5 to 2 mm (not significant) (cf. single /t/ and /t'/ in final position: 1 to 1.5 mm).

To summarize, the difference between the plain and palatalized coronals with respect to Tongue Body movement is significant over the VC interval on the vertical displacement (a higher Tongue Body raising for /t'/). The trajectory of /t/ also shows certain raising. The two consonants are similar in their horizontal Tongue Body movement.

5.3.3.6.2 Place

We have seen that both labials and coronals are significantly differentiated before consonants. In this section I focus on the place of articulation differences between palatalized (/p'/ vs. /t'/) and plain (/p/ vs. /t/) consonants.

5.3.3.6.2.1 Palatalized

Figure 5.14 presents the Tongue Body trajectories of palatalized stops /p'/ and /t'/ at two points in time for the three speakers. Again, the vertical displacement (TBy) is given in the left column and the horizontal displacement (TBx) is in the right column. For TBy higher values represent higher position of Tongue Body in the oral cavity. For TBx lower values correspond to a more front position of Tongue Body in the oral cavity. The overall Tongue Body differences between /p'/ and /t'/ at each point are given in Table 5.23. Non-significant differences (based on the corresponding main effects of Place and significant Place × Pal interactions) are marked with an asterisk.

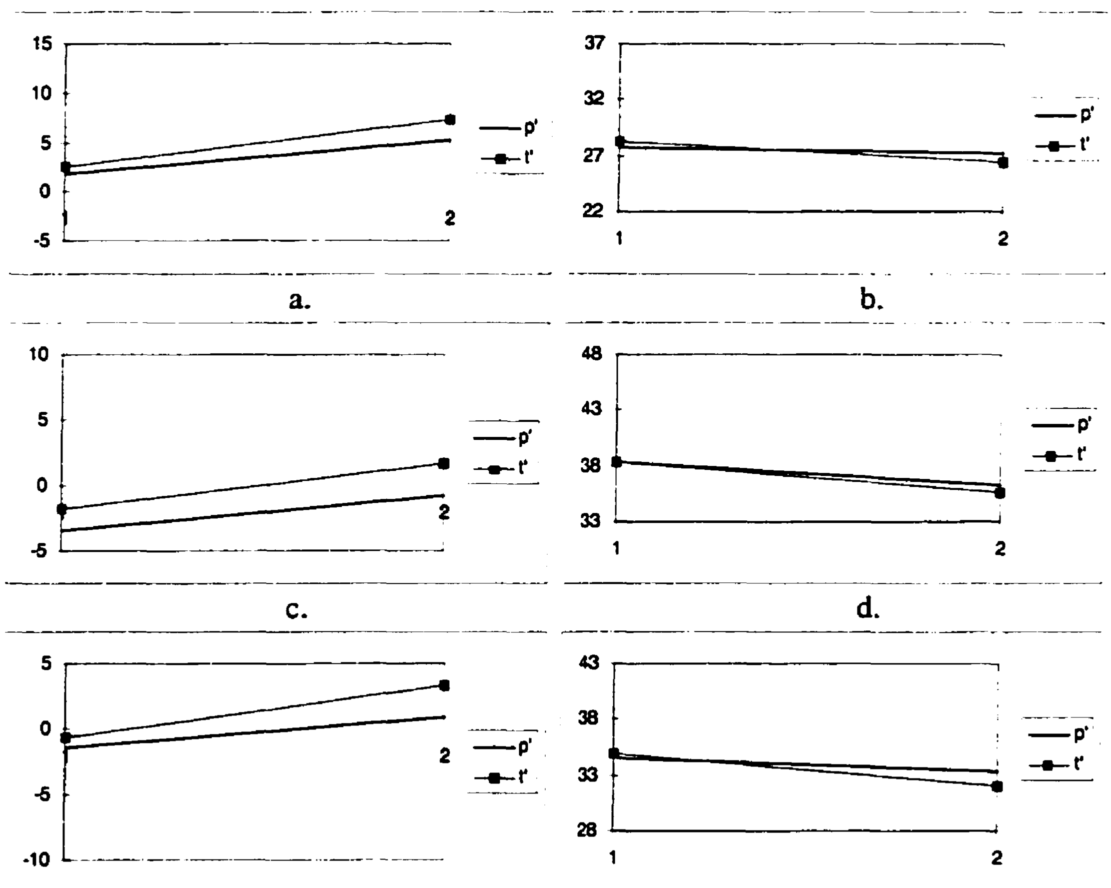


Figure 5.14 Movement of Tongue Body (vertical (ace) and horizontal (bdf)) during the articulation /p'/ and /t'/ as first consonant in a cluster, based on means at two points in time for Subjects 1 (ab), 2 (cd), and 3 (ed).

Table 5.23 Mean differences between /p'/ and /t'/

		VC1	VC2	Total
TBy	S1	1.64*	2.40	4.05
	S2	0.94*	2.00	2.94
	S3	0.98*	2.41	3.39
TBx	S1	0.62*	0.61*	1.23
	S2	0.60*	1.09*	1.69
	S3	0.84*	1.33	2.17

* Not significant difference ($p < .05$)

The articulations of both /p'/ and /t'/ involve Tongue Body raising, (Figure 5.14ace, the vertical displacement). For all the speakers, however, Tongue Body during the production of /t'/

is significantly higher (at VC2) than for /p'/. Recall that the same asymmetry between /p'/ and /t'/ was found for the single consonants (section 5.3.2.6.2.1). The total articulatory difference between the initial /p'/ and /t'/ in terms of TBy is from about 3 to 4 mm. (cf. single /p'/ and /t'/ in final position: 1.5 to 6 mm).

In terms of Tongue Body fronting, the two consonants are almost identical (except for a difference at VC2 for Subject 3), showing a high degree of Tongue Body fronting. The total articulatory difference between the initial /p'/ and /t'/ in terms of TBx is from about 1 to 2 mm (almost none significant) (cf. single /p'/ and /t'/ in final position: 2 to 3 mm).

In sum, in the preconsonantal environment the two palatalized consonants, /p'/ and /t'/, are very similar in their Tongue Body gestures (raising and fronting). The palatalized coronal, however, has a higher degree of Tongue Body raising.

5.3.3.6.2.2 Plain

Figure 5.15 presents the Tongue Body trajectories of plain stops /p/ and /t/ at two points in time for the three speakers. The overall Tongue Body differences between /p/ and /t/ at each point are given in Table 5.24.

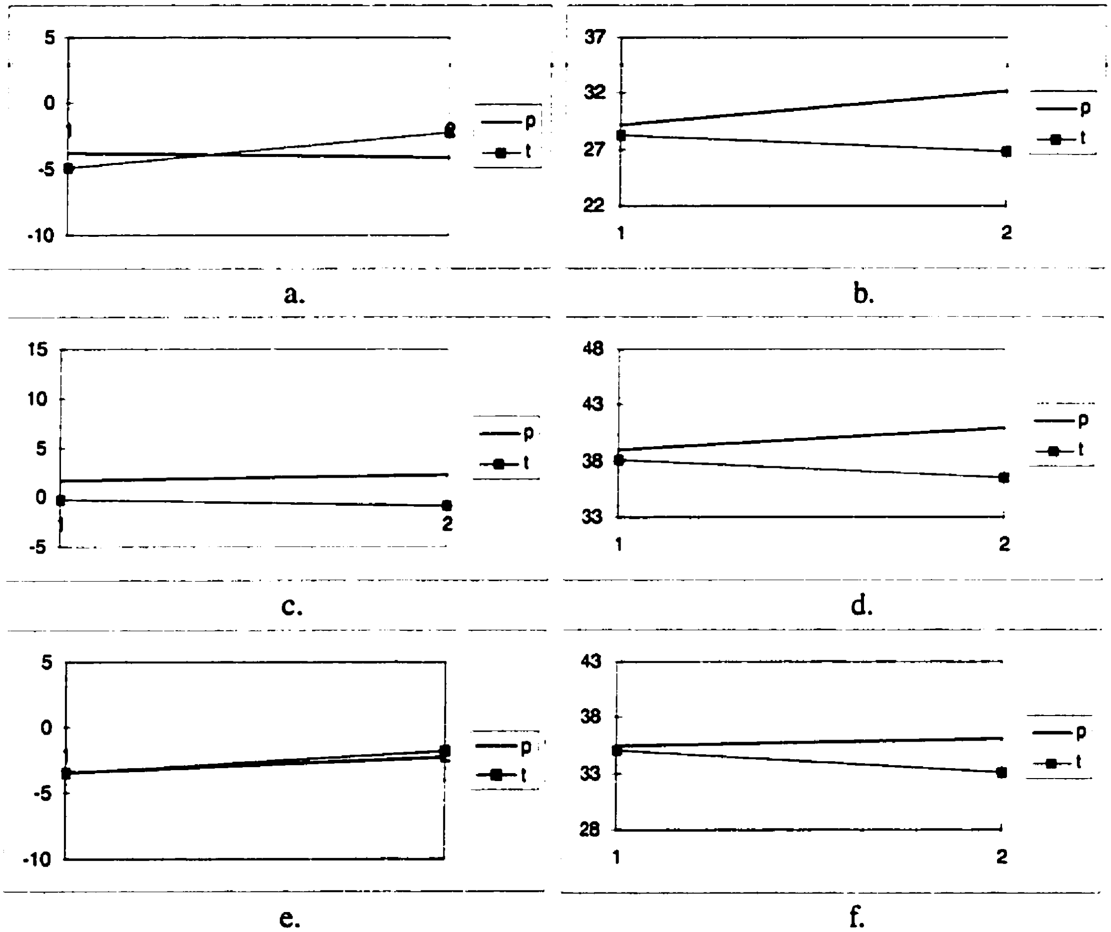


Figure 5.15 Movement of Tongue Body (vertical (ace) and horizontal (bdf)) during the articulation /p/ and /t/ as first consonant in a cluster, based on means at two points in time for Subjects 1 (ab), 2 (cd), and 3 (ed).

Table 5.24 Mean differences between /p/ and /t/

		VC1	VC2	Total
TBy	S1	1.09*	1.87	2.95
	S2	1.90	3.32	5.22
	S3	0.43*	1.23	1.66
TBx	S1	1.01	5.42	6.43
	S2	0.84*	4.37	5.21
	S3	0.81*	2.98	3.79

* Not significant difference (p < .05)

Although the two consonants are significantly different in TBy (at VC2 for all speakers and VC1 for Subject 2), these differences are relatively small. The total differences are from about 1.5 to 5 mm (cf. single /p/ and /t/ in final position: 2.5 to 5.5 mm).

It is the horizontal displacement, TBx, that most differentiates between /p/ and /t/ (at VC2 for all the speakers and also at VC1 for Subject 1). While the Tongue Body during /p/ is retracted, it is moved forward for /t/. The total articulatory difference between the preconsonantal /p/ and /t/ in terms of TBx is from about 4 to 6.5 mm. (cf. single /p'/ and /t'/ in final position: around 6 mm).

To summarize, labials and coronals (both plain and palatalized) in preconsonantal position are significantly different from each other in terms of Tongue Body magnitude. This difference is minor for palatalized /p'/ and /t'/. The plain consonants are differentiated more, primarily in the horizontal displacement, or in terms of tongue fronting/backing. It should be kept in mind that the major difference between labials and coronals is the primary gestures, Lip Aperture and Tongue Tip, discussed later in section 5.4.

Overall, we have seen that the Tongue Body position is the highest for /t'/, somewhat lower for /p'/, and then followed by /t/ and /p/. Tongue Body is the most fronted during the production of /p'/, /t'/, and /t/, and most backed for /p/. This is consistent with our results for single segments.

5.3.3.6.3 Environment

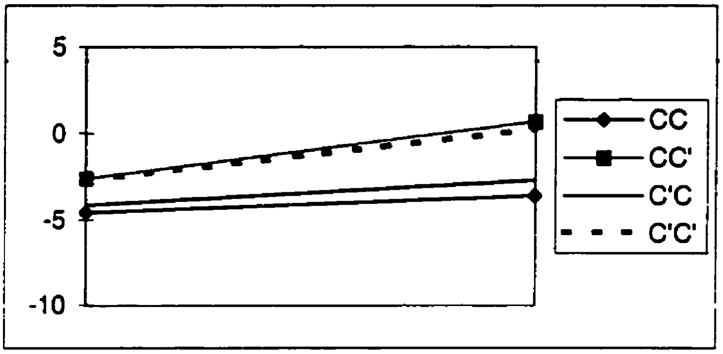
So far we have been examining the differences between consonants regardless of the quality of the following consonant. As we will see below, the quality of C_2 is an important factor in the production of plain and palatalized labials and coronals as the first segment of a cluster.

The statistical results (sections 5.3.3.2-5.3.3.4) showed that the VC transition trajectories contain information about whether C_2 in the cluster is plain or palatalized (the factor EnvPal). The overall differences in TBy movement between _C and _C' (regardless of C_1) are relatively small (about 2 mm at VC2 for both Subject 2 and Subject 3), but nevertheless, significant (except for Subject 1).[33] The Tongue Body gesture is higher in the environment before a palatalized consonant and lower when C_2 is plain. A similar effect was found for the horizontal movement, TBx: Tongue Body is more front in _C' and more backed before _C (Subjects 2 and 3).

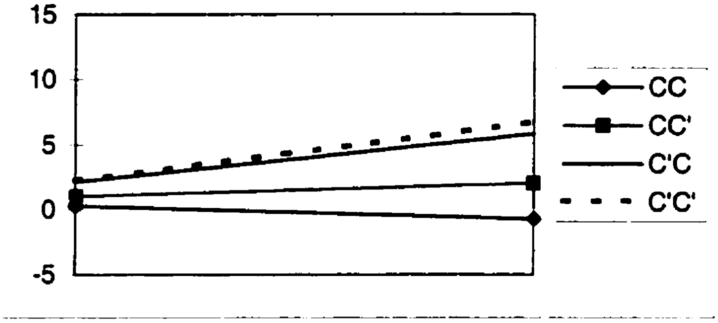
It is of particular interest to determine whether the articulation of plain and palatalized consonants is affected when the consonants are followed by segments that agree or disagree in terms of the secondary articulation. Here I refer to significant Pal × EnvPal interaction (significant across the speakers in TBy; Subject 1 and Subject 3 in TBx).

Figure 5.16 plots the TBy (vertical) trajectories before four combinations of plain and palatalized consonants (CC, CC', C'C, and C'C') for the three speakers.

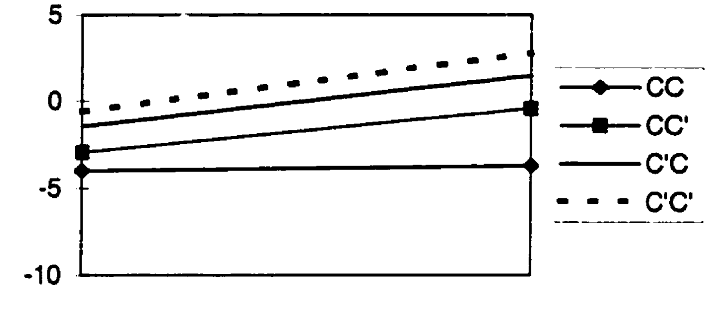
[33] The lack of the effect for Subject 1 can be attributed to the overall lower degree of overlap of stop gestures in clusters (see section 5.5.2.2).



a.



b.



c.

Figure 5.16 Movement of Tongue Body (vertical, TBy) during the articulation of plain and palatalized consonants followed by plain or palatalized segments, based on means at two points in time for Subjects 1 (a), 2 (b), and 3 (c).

The overall Tongue Body magnitude and slope during the transition both differ depending on the plain-palatalized quality of C_1 and C_2. For instance, the Tongue Body for a plain consonant is higher when it is followed by a palatalized segment (CC') than by a plain consonant (CC) (significant for Subject 2 and Subject 3 at VC2). At the same time, the Tongue Body magnitude of CC' is not identical to C'C': it is lower for Subject 2 and Subject 3, while slightly higher for Subject 1 (all significant at VC2). The trajectories for C'C and C'C' do not differ from each other (none are significant), while the Tongue Body movements prior to CC and C'C are different for Subject 1 (all significant at VC2). For all the subjects the overall difference between plain and palatalized consonants in TBy is higher before plain (CC vs. C'C: 1.3-8.5 mm) than palatalized segments (CC' vs. C'C': .5-6 mm).

Differences with respect to palatalization of C_1 and C_2 on the horizontal displacement, TBx, were found for Subject 3 only. In this case the Tongue Body during the articulation of a plain consonant was more front when C_2 was palatalized (CC'). This cluster did not significantly differ from C'C' in terms of Tongue Body fronting.

The significant Place × EnvPal interactions suggest that the effect of the following consonant (C or C') is different depending on the place of C_1. Thus, coronals are more sensitive to Environment (Tongue Body is raised in _C') than labials on vertical displacement, TBy (Subject 2, throughout the VC transition; see section 5.3.3.3). The reverse is true for horizontal displacement: labials are more affected (Tongue Body is fronted) by the following segment than coronals (whose Tongue Body is already in the front position) (Subjects 1 and 3, throughout the VC transition).

Although the place of articulation of the following consonant did not play a significant role (EnvPlace), it did show some influence when combined with secondary articulation of C_2 (EnvPal × EnvPlace). The effects, however, are not consistent among the speakers. Thus we find that a following palatalized coronal /t'/ induces fronting (Subject 1 at VC2) or raising (Subject 2 at VC2) of Tongue Body to a greater extent than a following /p'/. This is the opposite for Subject 3 (at VC1 and VC2): a following /p'/ induces more Tongue Body fronting of C_1 than a following /t'/ does.

The influence of a following labial is found for /p'/, whose Tongue Body gesture was less front in this environment (Place × Pal × EnvPlace, TBx Subject 1 at VC2), apparently due to the presence of the conflicting velarized gesture. The same interaction in TBy for Subject 2 points to the sensitivity of /t/ to the place of C_2: lowering of Tongue Body before labials and raising it before coronals.

In sum, whether the following consonant in the cluster is plain or palatalized is found to have a strong effect on the Tongue Body movement of the coda consonant during the VC transition. The plain consonant as C_1 is particularly affected if followed by a palatalized segment. This is due to the fact that the Tongue Body gesture for C_2 begins relatively early, thus affecting the transition to C_1. The opposite effect of Tongue Body backing or lowering is less robust. We also found certain asymmetries in terms of place and palatalization of C_1 and C_2.

5.3.3.7 Summary

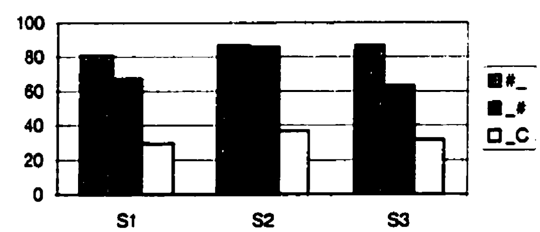
The results of the study of coda consonants followed by a consonant show that the magnitude of Tongue Body movement plays an important role in the articulation of all four stops. The Tongue Body contrasts are maintained in this position (as seen in their Tongue Body magnitude and slope differences); however, their magnitude is affected by the nature of the following consonant.

5.3.4 Single consonants vs. consonants in clusters

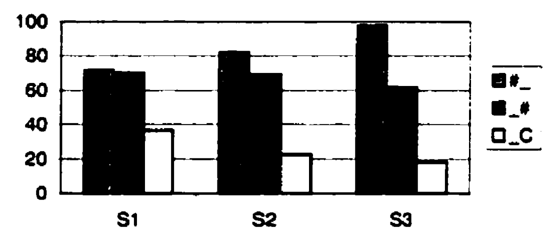
In this section I compare the Tongue Body results for single consonants (initial and final) with those in clusters.

Recall that utterances with single consonants had both VC transitions from the preceding vowel into the consonant and CV transitions into the following vowel. These transitions contained information about Tongue Body movement: its closing, constriction, and release, or absence of any movement. Unlike the case of single consonants, the first segment in a cluster, C_1, does not have the CV transition, since it is flanked by the closure of the second consonant. The information about the Tongue Body trajectory is made available only through the VC transition, the closing movement of Tongue Body.

We can compare single initial and final intervocalic consonants and segment C_1 in clusters in terms of overall Tongue Body differences (a sum of all differences at 4 points for single consonants and at 2 points for the segments in clusters; see section 5.3.1). Figure 5.17 shows these differences on the vertical (a) and horizontal (b) dimensions for the three speakers. It is obvious that while the overall Tongue Body differences between initial and final consonants are relatively small (and sometimes not significant), the preconsonantal environment shows a major drop in the overall contrast in terms of both TBy and TBx (from 60-100 mm to 20-40 mm).[34]



a.

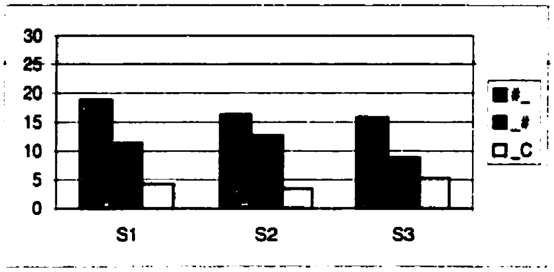


b.

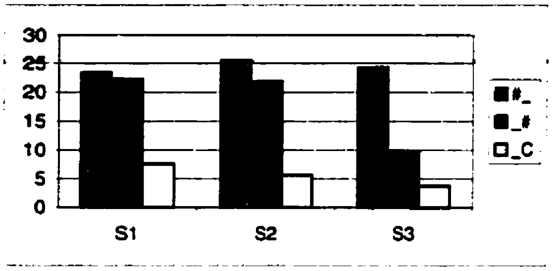
Figure 5.17 Overall Tongue Body (a. TBy and b. TBx) difference in three environments: initial (a#_a), final (a_#a), and preconsonantal (a_#C), for 3 Subjects

These differences are averaged over consonant and environment. It is also important to compare how the Tongue Body contrasts between consonants vary from position to position. Figure 5.18 presents the Tongue Body differences (TBy and TBx) in three environments for the contrasts /p/ vs. /p'/ and /t/ vs. /t'/.

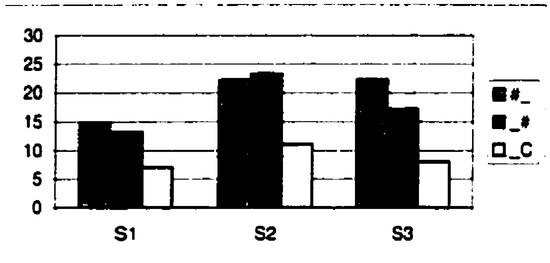
[34] The comparison of Tongue Body differences in Figure 5.17, however, does not take into account the fact that the differences continue to be substantial beyond the 30 ms interval investigated in this experiment. It also does not consider the relevance of burst quality in reflecting Tongue Body differences (see section 5.5).



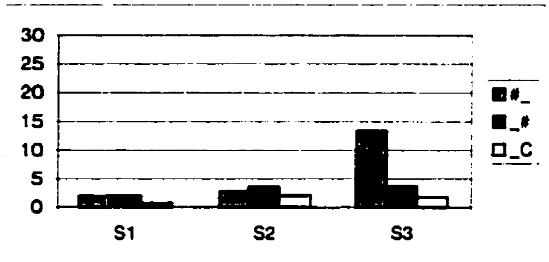
a. /p/



b. /p'/



c. /t/



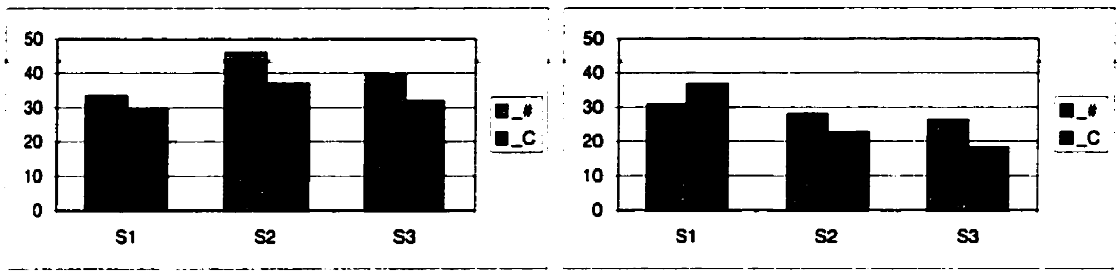
d. /t'/

Figure 5.18 Overall Tongue Body (ac. TBy and bd. TBx) differences /p/ vs. /p'/ (ab) and /t/ vs. /t'/ (cd) in three environments: initial (a#_a), final (a_#a), and preconsonantal (a_#C), for 3 Subjects

Recall that the palatalized labial /p'/ is distinguished from /p/ by both Tongue Body raising and Tongue Body fronting. The plain /p/ involves Tongue Body backing. The differences between /p/ and /p'/ are at their maximum during the CV transition. Thus, having only a VC transition leaves the Tongue Body contrast in plain and palatalized labials substantially reduced on both dimensions.

In the plain and palatalized coronals, which are differentiated by vertical displacement (TBy), the drop is also sizable. However, the lack of CV transition may not be as detrimental for them in the preconsonantal position, since the Tongue Body difference between the two is at its maximum during both VC and CV transitions.

Comparing the final and preconsonantal environments in terms of VC transition only, we notice (Figure 5.19) that there is a tendency (not always consistent) for fewer Tongue Body differences in the latter context. In general, however, the differences in the preconsonantal environment are still substantial (recall that the overall Tongue Body magnitude values at VC2 were not always significantly different between the two environments).



a. b.

Figure 5.19 Tongue Body (a. TBy and b. TBx) difference during VC transition in two environments, final (_#) and preconsonantal, (_C) for 3 Subjects

The Tongue Body differences for the preconsonantal position in Figure 5.19, however, are averaged for the following consonant. Recall that we found significant variation in terms of Tongue Body depending on the quality (particularly, plain or palatalized) of the following segment. Figure 5.20 illustrates this variability on the vertical displacement. The difference tends to be less before palatalized consonants as well as before homorganic segments (the latter difference, however, did not reach the significance level for all the subjects).

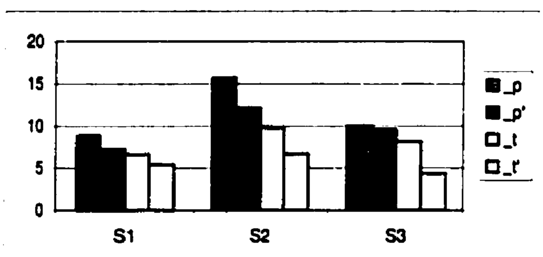


Figure 5.20 Tongue Body (TBy) difference during VC transition before four consonants for 3 Subjects

While the second consonant in the cluster, C_2, has not been systematically investigated in terms of its Tongue Body differences, informal observations lead me to suggest that its CV transition is similar to that of a single consonant. Unlike the latter, however, C_2 lacks the VC transition due to the closure of C_1. Thus, in this environment only the 'information' about Tongue Body release movement is directly available to the listener. Figure 5.21 shows the total differences in two environments (assuming that the postconsonantal position has the same CV transition as the initial single consonant).

Recall that the CV transition contains most of the information about the plain-palatalized contrast in labials and Tongue Body differences between the plain /p/ and /t/. The palatalized /p'/ and /t'/, however, are not significantly different from each other (at least at CV1).

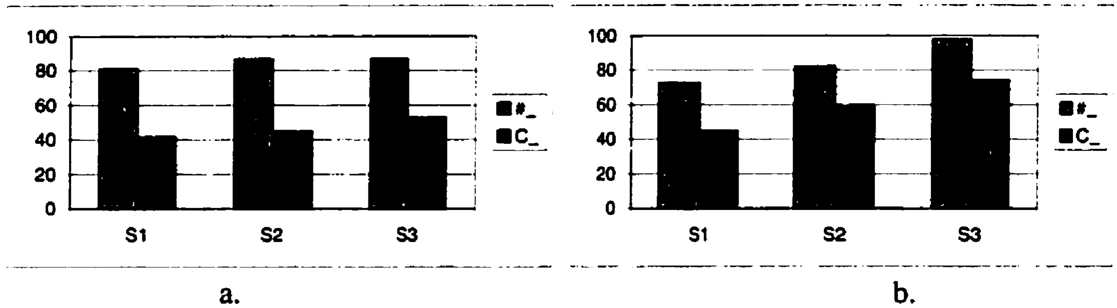


Figure 5.21 Tongue Body (a. TBy and b. TBx) differences during CV transition in two environments, initial and postconsonantal for 3 Subjects

In sum, unlike single consonants, the segments in clusters lack one of the transitions, and thus provide less information about the trajectories of the Tongue Body gesture. The first consonant of the cluster is at a particular disadvantage, due to both the overall drop in the Tongue Body difference and the influence of the following segment.

5.3.5 Summary and predictions

Based on the previous discussion of Tongue Body properties of the four stops under consideration and its variation over several contexts, I now formulate the observed asymmetries as statements that make explicit predictions about the recoverability of the plain and palatalized labials and coronals.

The results show that the contrast between the stops under consideration is best differentiated by Tongue Body movement in single consonants, particularly in initial (intervocalic) position. As stated in (5.1), we expect the contrast in initial position to be identified better than that in final position (5.1a). C_i and C_j stand for any of the four stops, /p/, /p'/, /t/, and /t'/, that differ in place or secondary articulation. The difference between these two environments, however, is not as great as between single consonants and segments in clusters (5.1b). In the cluster consonants, C_2 should be better recovered than C_1 (5.1c), since the 'information' about its Tongue Body opening movement (release) is directly available to a listener. And, finally, in preconsonantal position (C_1), higher correct identification is expected before the plain consonants than before palatalized ones (5.1d), because of the overall TB differences in contrast between the two contexts.

(5.1) Overall Tongue Body differences: Environments

a. C_i vs. C_j/(V)#_V > C_i vs. C_j/V_#(V)

b. C_i vs. C_j/(V)#_V or V_#(V) > C_i vs. C_j/(V)_C or C_(V)

c. C_i vs. C_j/ C_(V) > C_i vs. C_j/(V)_C

d. C_i vs. C_j/_C > C_i vs. C_j/_C'

Note that while all these statements apply to any contrast of the four consonants, they are particularly important for the plain-palatalized contrast. First, Tongue Body differences are more important in differentiating plain and palatalized segments than distinguishing the place contrast. Second, the plain and palatalized segments of each place share the same (or almost the same) primary articulator. Thus, Tongue Body movement is the only source of 'information' distinguishing the two consonants of the same place.

The statements in (5.1) may have a somewhat different effect depending on the place of articulation of the consonant. We saw that the contrast between plain and palatalized labials is more susceptible to environment than the coronal contrast, since the former shows less overall Tongue Body difference in final and, particularly, preconsonantal position. However, coronals appear to show more sensitivity to the nature of the following segment in preconsonantal position.

The two palatalized stops /p'/ and /t'/ are very similar in their Tongue Body fronting and raising gestures, showing some difference only during the VC transition. These two are likely candidates for confusion before consonants when no other information (primary place information available in the transitions and burst) is available.

The plain stops /p/ and /t/ differ in terms of Tongue Body mostly at the CV transition. Thus, when the place information is obscured, they are more likely to be confused in the preconsonantal position.

Looking at Tongue Body raising/lowering and fronting/backing, we can make a number of predictions about what consonants are most and least likely to be associated by a listener with high and front position of Tongue Body (in the absence of place information). Recall that the palatalized coronal /t'/ is characterized by the highest Tongue Body, followed by /p'/ with this gesture in a somewhat lower position. The plain /t/ and /p/ have lower Tongue Body, and do not differ from each other consistently in this respect. This is reflected in the hierarchy in (5.2).

(5.2) Degree of Tongue Body raising

t' > p' > t , p

The fronting hierarchy (5.3) groups together /t'/, /p'/, and /t/, which do not generally differ from each other in the degree of Tongue Body fronting. It juxtaposes these segments to /p/, which is characterized by the opposite movement of Tongue Body (backing).

(5.3) Degree of Tongue Body fronting

t', p', t > p

Having identified a high Tongue Body position in a stimulus, a listener is expected to identify it primarily with /t'/, followed by /p'/ (in the absence of place information). It is less likely that his/her response would be /t/ or /p/. The recovered front position of the Tongue Body gesture can be interpreted as one of the three consonants /t'/, /p'/ or /t/, with /p/ as the least likely response. It is obvious that the two scales agree in differentiating /t'/ and /p'/ on the one hand

and /p/ on the other. At the same time, they make conflicting predictions about the plain coronal /t/.

Note that the relative height of Tongue Body of /p'/ is somewhat lower in final and preconsonantal positions, thus moving it closer to /p/. We would expect that in these environments /p'/ has a slightly higher chance of being perceived as plain.

A lower position of Tongue Body during the articulation of /t'/ before plain homorganic segments may also be interpreted as a lack of raising of this gesture. On the other hand, the anticipatory raising and fronting movement of Tongue Body towards the palatalized C_2 also raises and fronts Tongue Body of a plain C_1 (/p/ or /t/), and thus increases the likelihood of responses in favour of palatalized /p'/ and /t'/.

These predictions about the perception of relative differences, raising and fronting of Tongue Body, and its modification across environments do not take into account the primary articulator (see section 5.4) and the acoustic burst differences (to be discussed in section 5.5). Taken together all the articulatory and acoustic properties will be tested in the perceptual experiment in Chapter 6.

5.4 Primary gesture (single consonants)

The goal of this experiment is to determine whether there are any articulatory differences between plain and palatalized consonants in terms of their primary articulators, Lip Aperture (LA) and Tongue Tip (TT). I examine single consonants in two environments, initial and final.

5.4.1 Analysis

The measurements of the primary gestures were made at the peaks (maximum values of constriction) of their trajectories, determined by a Mavis procedure (see section 5.2.3). For Tongue Tip the measurements for both vertical (TTy) and horizontal (TTx) values were made at the maximum of TTy, since the two were almost simultaneous. The same was done for the Upper

Lip (ULy and ULx) articulator for Subject 1. For Subjects 2 and 3 the minimum tangential value of Lip Aperture, LAxy (combined Upper and Lower Lip values) were measured.

I analyzed the parameters of interest in separate ANOVAs for each primary articulator, Lip Aperture and Tongue Tip, for each subject. Each test involved two between-items factors, Palatalization (*Pal:* plain and palatalized), and Environment (*Env*: initial and final) and one dependent variable: Lip Aperture (LAxy) magnitude (or ULy and LLx magnitude for Subject 1), and Tongue Tip (TTy and TTx) magnitude. Tukey HSD post-hoc tests were performed to investigate significant interactions. We are particularly interested in finding out if there is a main effect of Pal, which would mean that plain and palatalized consonants of the same place of articulation differ with respect to the primary articulation, Lips or Tongue Tip.

5.4.2 Overall results

In this section I give a general overview of the results. The results by major groups of stimuli as well as their discussion are presented in more detail in later sections. Tables 5.25-5.27 present means and standard deviations for the magnitude of primary gestures, Lip Aperture and Tongue Tip, during the articulation of initial and final /p/, /p'/, /t/, and /t'/ respectively. The results are given by subject (based on the average per 5 tokens per subject) and subdivided by vertical (TBy) and horizontal (TBx) displacement (where appropriate). The consonants of interest in each utterance are given in bold. Means for each consonant in the two environments are also presented. Note that the absolute values cannot be compared between subjects, since they are based on an arbitrary starting point due to the individual differences in the speakers' articulatory apparatuses.

Table 5.25 Magnitude of Upper Lip (UL) and Tongue Tip (TT): Subject 1

	ULy		ULx				TTy		TTx	
	mean	SD	mean	SD			mean	SD	mean	SD
ta papy	-16.36	1.06	-18.54	.87		ta tapy	-3.46	.55	4.22	.81
tap apy	-15.42	1.15	-18.62	.97		tat apy	-3.4	.47	4.5	.58
mean	*-15.89*		*-18.58*			*mean*	*-3.43*		*4.36*	
ta p'apy	-15.96	1.18	-18.56	1.04		ta t'apy	-3.54	.44	3.64	.73
tap' apy	-16.04	1.13	-18.36	1.16		tat' apy	-3.40	.44	4.12	.57
mean	*-16*		*-18.46*			*mean*	*-3.47*		*3.88*	

Table 5.26 Magnitude of Lip Aperture (LA) and Tongue Tip (TT): Subject 2

	LAxy				TTy		TTx	
	mean	SD			mean	SD	mean	SD
ta papy	2.68	1.5		ta tapy	4.00	1.42	10.94	1.67
tap apy	2.76	1.05		tat apy	3.96	1.79	11.18	2.35
mean	*2.72*			*mean*	*3.98*		*11.06*	
ta p'apy	2.5	.94		ta t'apy	6.82	2.19	14.42	1.59
tap' apy	3.04	1.19		tat' apy	6.22	2.69	13.04	2.61
mean	*2.77*			*mean*	*6.52*		*13.73*	

Table 5.27 Magnitude of Lip Aperture (LA) and Tongue Tip (TT): Subject 3

	LAxy				TTy		TTx	
	mean	SD			mean	SD	mean	SD
ta papy	19.25	1.70		ta tapy	-4.35	.59	13.28	1.83
tap apy	16.77	.94		tat apy	-4.34	.69	12.20	1.67
mean	*18.01*			*mean*	*-4.35*		*12.74*	
ta p'apy	18.18	.82		ta t'apy	-1.13	1.44	11.06	1.51
tap' apy	17.14	.97		tat' apy	-1.17	.95	13.08	1.74
mean	*17.66*			*mean*	*-1.15*		*12.07*	

Consider how to interpret these tables, taking the right column of Table 5.25 as an example. During the production of the second plain coronal /t/ of *ta tapy* the Tongue Tip gesture of Subject 1 was at -3.46 (.55) on the vertical dimension and at 4.22 (.81) on the horizontal dimension. This is the maximum point of its constriction (at the teeth). Note that the first number here stands for the mean value and the second number is the standard deviation. Values given in the discussion that follows are for the final /t/ (in *tat apy*) and the palatalized /t'/ in two environments.

Further I present the statistical results by articulator and by subject (sections 5.4.3.1-5.4.3.3). These are followed by the discussion (5.4.3.4).

5.4.3 Results: Lip Aperture

5.4.3.1 Subject 1 (statistics)

Table 5.28 shows the ANOVA results for the Upper Lip magnitude (the horizontal, ULy, and vertical, ULx displacement) for single initial and final labials, as pronounced by Subject 1.

Table 5.28 Magnitude of Upper Lip: Subject 1 (statistics)

	ULy		ULx	
	F	*p*	*F*	*p*
Pal	.05	.830	.07	.795
Env	.72	.410	.02	.896
Pal x Env	1.01	.330	.10	.762

As we can see, none of the factors was significant (i.e. $p < .05$) either at ULy (Pal [$F(1,20) = .05$, $p = .83$], Env [$F(1,20) = .72$, $p = .41$]) or at ULx (Pal [$F(1,20) = .07$, $p = .795$], Env [$F(1,20) = .02$, $p = .896$]). There was no significant interaction of Pal × Env ([$F(1,20) = .10$, $p = .762$]).

5.4.3.2 Subject 2 (statistics)

Table 5.29 gives the ANOVA results for the Lip Aperture magnitude (combined ULy and ULx displacement) for Subject 2. Again, there was no main effect of neither of Pal nor of Env. There were no significant interactions.

Table 5.29 Magnitude of Lip Aperture: Subject 2 (statistics)

	LAxy	
	F	*p*
Pal	.01	.926
Env	.34	.570
Pal x Env	.19	.670

5.4.3.3 Subject 3 (statistics)

Table 5.30 gives the statistics results for the Lip Aperture for Subject 3. The significant results are given in bold.

Table 5.30 Magnitude of Lip Aperture: Subject 3 (statistics)

	LAxy	
	F	*p*
Pal	.45	.512
Env	**11.59**	**.004**
Pal x Env	1.93	.184

Here we find a main effect of Env [F(1,20) = 11.59, p = .004]. Pal was not significant. There were no significant interactions.

The comparison of mean values for initial and final labials shows that final labials have lower magnitude than initial ones.

5.4.3.4 Discussion

The results show that plain and palatalized labials do not differ in their primary articulator, Lip Aperture. In other words, the addition of the secondary articulation does not affect the magnitude of the primary gesture. Env was found to be a factor for only one subject. In this case final labials show some reduction in magnitude. The reduction of gestures in the coda position is a common process found across languages (see Gick 1999: 58 and references therein).

Now I turn to the coronal plain and palatalized consonants and examine the results of Tongue Tip magnitude. The statistical results by articulator and by subject are given in sections 5.4.4.1-5.4.4.3). These are followed by the discussion (5.4.4.4).

5.4.4 Results: Tongue Tip

5.4.4.1 Subject 1 (statistics)

Table 5.31 shows the ANOVA results for the Tongue Tip (TT) magnitude (the horizontal, TTy, and vertical, TTx displacement) for single initial and final coronals, as pronounced by Subject 1.

Table 5.31 Tongue Tip: Subject 1 (statistics)

	TTy		TTx	
	F	*p*	*F*	*p*
Pal	.04	.854	2.49	.134
Env	.22	.646	1.56	.230
Pal × Env	.04	.854	.11	.747

For this subject none of the factors were significant either at TTy or at TTx. There was no significant interaction of Pal × Env.

5.4.4.2 Subject 2 (statistics)

Table 5.32 gives the Tongue Tip results for Subject 2.

Table 5.32 Tongue Tip: Subject 1 (statistics)

	TTy		TTx	
	F	*p*	*F*	*p*
Pal	**7.48**	**.015**	**8.08**	**.012**
Env	.12	.735	.37	.553
Pal x Env	.09	.767	.74	.401

There was a main effect of Pal for both displacements, TTy [F(1,20) = 7.48, p = .015] and TTx [F(1,20) = 8.08, p = .012]. Env was not significant. No significant interactions were found. (Means: TTy /t/ 3.98, /t'/ 6.52; TTx /t/ 11.06, /t'/ 13.73).

The comparison mean values shows that the values for palatalized coronals are lower and more back than for their plain counterparts.

5.4.4.3 Subject 3 (statistics)

Table 5.33 gives the Tongue Tip statistics results for Subject 3.

Table 5.33 Tongue Tip: Subject 1 (statistics)

	TTy		TTx	
	F	*p*	*F*	*p*
Pal	**53.46**	**.000**	.91	.355
Env	.00	.969	.30	.591
Pal x Env	.00	.955	**4.46**	**.051**

Here we see a highly significant main effect of Pal on the vertical displacement, TTy [F(1,20) = 53.46, p = .000] (Means: /t/ -4.35, /t'/ -1.15). Pal on the horizontal displacement, TTx, and Env were not significant. The Pal × Env interaction at TTx was slightly above the significance level [F(1,20) = 4.46, p = .051]. (Means: #t 13.28, t# 12.20, #t' 11.06, t'# 13.08)

5.4.4.4 Discussion

The results for Tongue Tip magnitude show variability among the speakers. While /t/ and /t'/ have the same magnitude for Subject 1, these two segments are different from each other in terms of Tongue Tip magnitude, at least on some parameters, for Subject 2 and Subject 3. During the production of the palatalized /t'/ by these speakers Tongue Tip is always higher than for the plain /t/. The difference between /t/ and /t'/ for both subjects, however, is fairly small, about 2.5-3 mm. In addition, Subject 2 has the Tongue Tip in a slightly more back position (by 2.5 mm).[35]

Given that no palate traces were made in this experiment, it is hard to determine the exact constriction location for these consonants. However, the higher values on vertical displacement, found for Subjects 2 and 3, show that, unlike for /t/, the constriction for /t'/ is made at the alveolar ridge rather than at the upper teeth (or at least farther back on the teeth-alveolar ridge continuum). The differences on the horizontal displacement (Subject 2) may be due to whether the tip or the blade (lamina) of the tongue is used.[36]

The variation found between speakers may be attributed to individual differences, dialect differences (recall that Subjects 2 and 3 speak the same variety of Standard Russian), or a combination of the two. It should be noted, however, that the articulatory differences between /t/ and /t'/ in the position of Tongue Tip have been previously documented for Standard Russian, including its Moscow variety (Bolla 1981, Skalozub 1963). While the plain /t/ is often described as laminal dental, the palatalized /t'/ is treated as laminal alveolar.

These relatively small differences in the articulation of /t'/ are unlikely to have noticeable acoustic consequences for VC and CV transitions. However, they may contribute to the difference between the bursts of the two coronal stops (and possibly between the subjects) (see section 5.5).

5.4.5 Summary and conclusion

The results of the experiment reveal that single labial plain and palatalized consonants do not differ in terms of their primary place of articulation. The results for the coronal stops are less

[35] For Subject 3 the horizontal dimension is marginally relevant to initial consonants, where the articulation of /t'/ is more front than that of /t/ (which is the opposite of what we find for Subject 2).

[36] Recall from section 5.1 that both Russian /t/ and /t'/ are commonly described as laminal, i.e. articulated with the tongue front (e.g. Bolla 1981). It should be noted that for all the subjects the receivers were placed not exactly on the tip of the tongue proper (which is difficult for technical reasons), but on the upper surface of the tongue close to the tip, i.e. approximately between the tip and the blade.

consistent: while the Tongue Tip articulation of /t/ and /t'/ is the same for Subject 1, it is to various degrees different for Subjects 2 and 3. In these cases Tongue Tip of the palatalized coronal has a slightly higher position than that of its plain counterpart.

The explanation for the tendency to move the primary articulation for /t'/ further back within its range is in the fact that the articulators involved in the production of /t'/ (Tongue Tip and Tongue Body) are coupled such that the movement of one is likely to affect the movement of the other (see the discussion in section 5.3.2.6.1.2).

The primary articulator may be affected by position in the word, but the reduction of final gestures is not consistent across subjects and is limited to only some consonants.

In general, we can conclude that the primary articulator plays a very minor role in distinguishing the plain-palatalized contrast (for coronals), or no role at all (for labials). This makes further investigation of the magnitude of primary place gestures in consonant clusters unnecessary.[37] The two gestures, however, are crucial in distinguishing the primary place of articulaton: labial vs. coronal. I next turn to the analysis of articulatory release and acoustic burst.

5.5 Articulatory release and acoustic burst

5.5.0 Introduction

In the previous section I examined the Tongue Body trajectories during the articulation of plain and palatalized labial and coronal stops. The acoustic consequences of these movements (and the primary gesture movements) are available to the listener in the VC and CV transitions. Recall from Chapter 3 that these transitions are important in the perception of the plain-palatalized contrast. In addition to the transitions, the information that serves to identify /p/, /p'/, /t/, and /t'/ is encoded in the release burst of the stops, a period of noise that follows the silent closure. In this section I examine the properties of the articulatory release that gives rise to the acoustic burst. I first discuss release in single stops and then in consonant clusters, both hetero-organic and homorganic. I show the crucial importance of overlap of gestures in the presence or absence of an acoustic burst.

It is beyond the scope of this section to investigate the spectral acoustic properties of bursts (see the acoustic study of /t/ and /t'/ in Chapter 3; also Bolla 1981, Bondarko 1981)).

5.5.1 Single stops

As in the case of transitions, burst frequencies reflect the primary and secondary gestures involved in an articulation. Figure 5.22 shows articulatory gestures during the production of the four stops, /p/, /p'/, /t/, and /t'/, and the resulting acoustic signal for Subject 2. The two lower images present the gesture trajectories (Tongue Body and Tongue Tip or Lip Aperture). Lip Aperture combines both vertical and horizontal dimensions (xy), while only vertical movement (y) is given for Tongue Tip and Tongue Body. The two upper images give the acoustic signal for both the entire utterance and the interval of interest.

[37] The mean TTy and TTx values of coronal tokens in clusters /tp/ and /t'p/ are almost the same as for the single consonants. It is often technically impossible to determine a peak of the gesture of C1 if it is followed by a homorganic consonants, since we tend to find one long gesture shared by the two consonants.

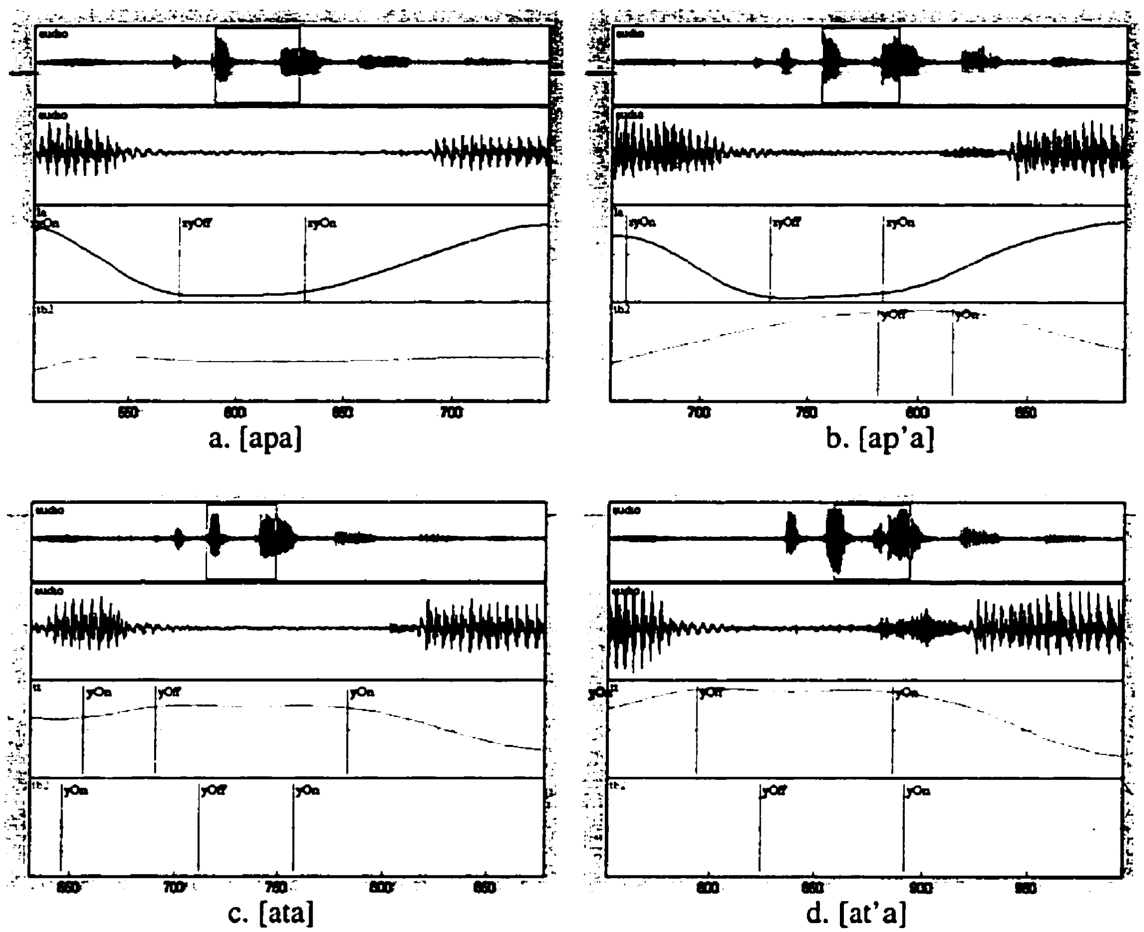


Figure 5.22 Articulatory gestures and acoustic burst of [p] (a), [p'] (b), [t] (c), and [t'] (d), as pronounced by Subject 2 (tokens n4_27, n3_54, n4_37, n4_44)

The labels on the trajectories mark the beginning of movement (xyOn or yOn), the achievement of the articulatory target or constriction (xyOff or yOff), and the beginning of the articulatory release (the second set of xyOn or yOn). Since there is no complete constriction during the closing movement of the primary gesture (Lip Aperture or Tongue Tip), the air is coming through the oral cavity, and we can observe the high energy of the vowel in the acoustic signal (which is also manifested in the spectral formants during the VC transition, not shown here). During the stop constriction the passage is sealed off, and no acoustic energy is available (acoustic closure, or silence). At the beginning of the articulatory release of Lip Aperture or Tongue Tip the accumulated air bursts through the narrow constriction and results in an interval of acoustic noise, a random energy. The duration of this noise depends on how long the narrow constriction is maintained at the release. The quality (or spectral characteristics) of the noise is dependent on the articulators involved as well as the shape of these articulators.

We can see from Figure 5.22 that there are crucial differences between the four stops. The burst is relatively short and weak for labials (ab) due to the relatively abrupt movement of Lips and the shape of the oral cavity, which creates weak noise at low frequencies. The narrow (glide-like) constriction of fronted and raised Tongue Body at the hard palate creates an additional source of noise for /p'/. Due to this, the release burst of the palatalized labial is longer and has stronger energy than that of /p/. Notice that the peak of the Tongue Body gesture is timed with the end of constriction of Lip Aperture, or the beginning of its release. Unlike for /p'/, the backed Tongue Body apparently does not create a narrow enough constriction with the velum to be able to generate additional noise.

In general, the noise period is longer for coronals (cd). This is due to the nature of the Tongue Tip constriction. The laminal shape of the tongue, especially for /t'/, leads to greater contact with the upper teeth and the alveolar ridge and thus to a slower articulatory release. The noise at mid and higher frequencies results from the shape of the oral cavity during the production of coronal stops. The raising of the Tongue Body for /t'/ creates an additional narrow constriction at the palate that greatly modifies the spectrum by adding strong high frequency noise.

The diagrams in Figure 5.22 show word-initial consonants. An informal examination of burst quality in word-final single consonants followed by vowel /a/ did not reveal any substantial differences from the word-initial position.

In sum, the differences in articulation of the four stops under consideration result in different qualities of release burst. The bursts of palatalized stops are acoustically more salient than those of the corresponding plain stops, being characterized by longer duration and higher energy at high frequencies. At the same time, the coronal plain and palatalized stops have more salient (longer duration and higher energy) bursts than the corresponding labials.

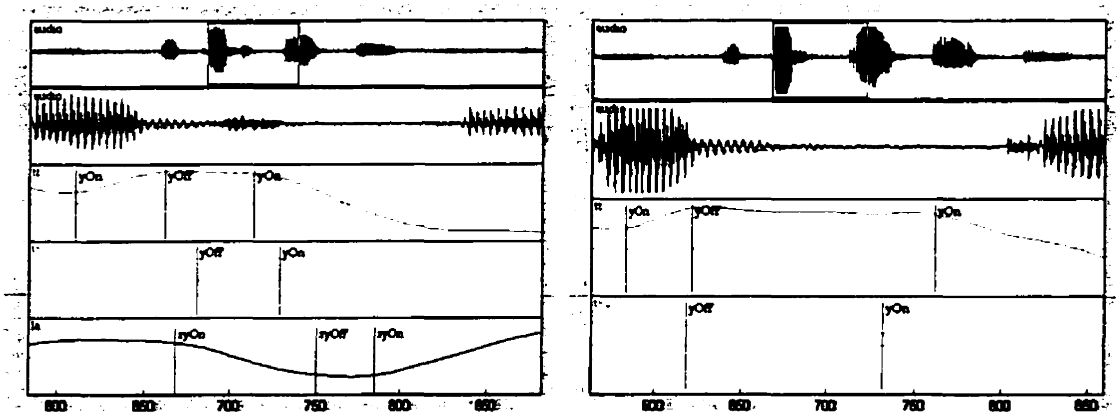
The presence of the burst prior to the following vowel makes the right edge of the stop crucially important, since it provides additional information about the quality of the consonant, its place and its secondary articulation. One might expect then that the burst should play a special role in distinguishing two stops with similar Tongue Body trajectories. For instance, /p'/ and /t'/ do not differ in their Tongue Body gestures during the CV transition. However, they do differ in the quality of their burst: a weaker noise with some low energy at high frequencies for /p'/ and a strong strident noise at higher frequencies for /t'/. The duration of the two release bursts may also be important in differentiating the two: a relatively short burst of /p'/ vs. a very long noise element of /t'/, which can be compared to a period of fricative noise of an affricate.

In the next sections I examine how the properties of release (and its consequences) are modified in clusters.

5.5.2 Stops in clusters

What happens when a stop is followed by another consonant? In Chapter 3 we saw that the likelihood of burst release of /t/ and /t'/ depend on the place of articulation of the following consonant. The stops are audibly released significantly more often before hetero-organic /k/ than before homorganic /n/ and /s/. That acoustic study did not investigate clusters with labials as C_1 (i.e. /pt/, /pp/, /p't/, /pp/) and as C_2 (/tp/ and /t'p/. It also did not analyze stops as the second consonants in clusters (C_2). However, the findings for clusters of coronal stop plus /k/, /n/, and /s/ were predicted to hold for all hetero- and homorganic sequences.

A preliminary analysis of the articulatory data shows that these predictions are largely supported by the new data. Figure 5.23 presents the articulatory gestures and their acoustic consequences during the production of hetero-organic cluster /t'p/ (*in tat' papy*) and homorganic sequence (*tat' tapy*) as pronounced by Subject 2. Notice that the release of the first gesture, Tongue Tip, begins before the second gesture, Lip Aperture, achieves its target. The distance between the two constrictions will be referred to as a 'lag' between two gestures, which is, in fact, the opposite of gestural overlap. This timing of articulatory gestures results in an acoustic burst before the second closure. Also notice that the palatalized gesture here is timed with the beginning of Tongue Tip release, or, in acoustic terms, with the burst.



a. t'p b. t't

Figure 5.23 Gestures and acoustic signals during the production of the clusters /t'p/ (in *tat' papy*) and /t't/ (in *tat' tapy*), Subject 2 (tokens n4_17 and n3_65)

Contrary to this presence of burst in the hetero-organic cluster (Figure 5.23a), there is only one constriction (Tongue Tip) during the production of the homorganic sequence /t't/ (Figure 5.23b). The Tongue Tip gesture achieves its target for the first consonant, /t'/, and stays there for the production of the second /t/ (see Figure 5.1). As a result, no air leaves the oral cavity throughout the closure, and consequently, there is no burst in the acoustic signal. Notice that the duration of the constriction (and the acoustic closure) is about twice as long for a single /t/. It should be noted that it is still possible to have a burst in a homorganic cluster (see section 5.5.2.4); however, it comes at the cost of additional Tongue Tip movements (release of the first gesture and achievement of the second gesture). The same holds for hetero-organic and homorganic sequences with C_1 as a labial.

Below I examine the question of presence or absence of burst in more detail, in hetero-organic and homorganic clusters.

5.5.2.1 Hetero-organic clusters

I have shown above (and in Chapter 3) that the first gesture in a hetero-organic cluster in Russian is frequently released. Is timing the same for all the stop clusters of interest (/pt/, /pt'/, /p't/, /p't'/, /tp/, tp'/, /t'p/, and /t'p'/)? Are there any differences depending on the place and secondary articulation of C_1 and C_2?

5.5.2.1.1 Overview

A preliminary analysis determined that there are some timing differences between clusters. Compare two tokens, /tp/ and /pt/, pronounced by Subject 3, presented in Figure 5.24. They involve the same two constrictions of gestures, Lip Aperture and Tongue Tip, although in the opposite temporal order. The constriction of Tongue Tip is followed by the 'plateau' of Lip Aperture in the case of /tp/, and Lip Aperture is followed by Tongue Tip in the production of /pt/. It is important to notice the difference in timing of the two gestures: while there is a gap of about 25 ms between the two constrictions in /tp/, the two constrictions in /pt/ overlap: the Tongue Tip target is achieved prior to the release of Lip Aperture. The consequences of these two timing patterns are: an acoustic burst between the two closures in /tp/ and no burst in /pt/.

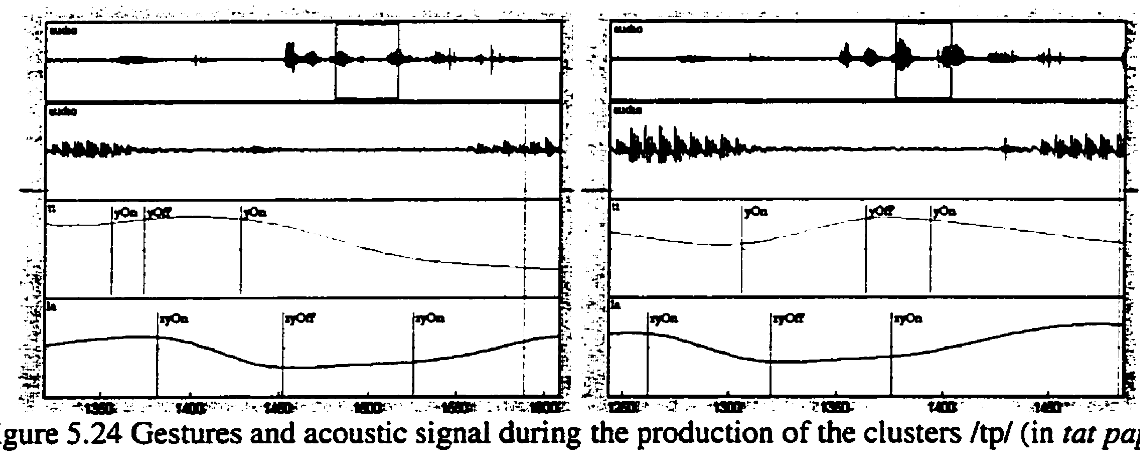


Figure 5.24 Gestures and acoustic signal during the production of the clusters /tp/ (in *tat papy*) and /pt/ (in *tap tapy*), Subject 3 (tokens n3_63 and n1_03)

5.5.2.1.2 Analysis

To determine the nature of these asymmetries I examined all tokens with hetero-organic labial and coronal clusters for the three subjects, for a total of 120 tokens: 8 utterances × 5 repetitions × 3 subjects). I used a measure of *lag*, a period of time (in ms) between the release of the first gesture and the achievement of articulatory target by the second gesture (Kochetov & Goldstein 2001). This is illustrated in Figure 5.25, where the lag is the distance between the beginning of release of the Tongue Tip gesture and the achievement of the target by Lip Aperture. This lag has a positive value, about 50 ms, and it results in an acoustic burst. The lag can be negative if the two constrictions overlap. This will result in a long stop closure, uninterrupted by burst.

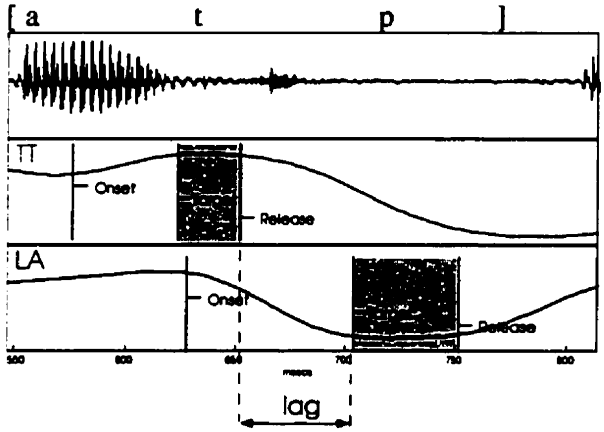


Figure 5.25 Gestures and acoustic signal during the production of the clusters /tp/ (in *tat papy*), Subject 2

In addition, I compared the clusters in terms of their *release rate*, the proportion of tokens for a particular cluster, where C_1 was released. The first consonant was considered to be released if the lag value was positive, and unreleased if it was negative. If all the tokens for a given cluster were released, the sequence was assigned the rate of 1.00. If none of them were released, the rate for the cluster was .00. The two measures (lag and release rate) are similar, but they do not necessarily replicate each other, since, as we will see, a cluster may have an overall negative lag, while having C_1 released in some tokens (positive lag). The measure of release rate will later allow us to compare hetero-organic clusters to the homorganic ones (section 5.5.2.4.2).

5.5.2.1.3 Results and discussion

Means and standard deviations (based on 5 tokens) for lag in these clusters are given in Table 5.34 and those for release rate are presented in Table 5.35. The results are organized by the order of place of articulation of the stops in a cluster (labial-coronal and coronal-labial) and by speaker. The means by the order of place and the overall means per subjects are also calculated.

Recall that positive lag values mean that there was a certain distance between the constrictions of C_1 and C_2 (no overlap). The negative numbers represent the fact that the two gestures were overlapped, i.e. the constriction for the second consonant was made before the first constriction was released. In acoustic terms there is a release burst of the first consonant in a cluster with positive lag and no release of this consonant in a sequence with negative lag overlap.

For example, the mean lag for /pt/ for Subject 3 is -14.8 (16.6) ms. This means that the gesture of Lip Aperture was overlapped with the following gesture of Tongue Tip by 14.8 ms. At the same, the cluster with the same consonants, but in the reverse order, /tp/, shows a positive lag of 5.0 (19.49) ms, i.e. the constriction of Tongue Tip and Lip Aperture were apart by 5.0 ms. This will result in the absence of release burst of /p/ in the cluster /pt/ and the presence of it, although a short one, in the cluster /tp/.

Table 5.34 The degree of lag in hetero-organic clusters: Means and standard deviations

	Subject 1		Subject 2		Subject 3	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*
pt	28.8	4.38	2.5	19.07	-14.8	16.16
pt'	20.8	20.72	11.2	21.29	-18.8	18.36
p't	42.4	13.37	1.2	18.69	-17.3	20.82
p't'	38	13.64	8.4	10.90	-32.4	38.82
mean	*32.5*		*5.83*		*-20.8*	
tp	53.6	11.26	31.2	15.85	5.0	19.49
tp'	49.0	23.86	2.5	47.31	-11.5	27.20
t'p	70.0	11.58	35.2	10.45	16.0	5.66
t'p'	50.8	6.10	4.0	22.93	20.8	12.13
mean	*55.85*		*18.23*		*7.58*	
overall mean	**44.18**		**12.03**		**-6.63**	

Table 5.35 Release rate in hetero-organic clusters: Means

	Subject 1	Subject 2	Subject 3
pt	1.00	.60	.20
pt'	.80	.40	.20
p't	1.00	.40	.33
p't'	1.00	.60	.20
mean	*0.95*	*0.50*	*0.23*
tp	1.00	1.00	.50
tp'	1.00	.80	.25
t'p	1.00	1.00	1.00
t'p'	1.00	.40	1.00
mean	*1.00*	*0.80*	*0.69*
overall mean	**0.98**	**0.65**	**0.54**

The release rate of .60 for the cluster /pt/ as pronounced by Subject 2 means that 60% of the tokens with this cluster were released, while 40% overlapped. Note that this variation is not reflected in the overall lag for this cluster, which is positive (2.5 ms). The cluster /tp/, on the other hand, shows a rate of 1.00, which indicates that 100% of the tokens were released (cf. the lag of 15.85 ms). It is important to note that while many of the clusters pronounced by Subject 3 have a negative lag, they always have a release rate higher than .00, i.e. C_1 is released in at least one token per cluster.

Notice that speakers vary both in the relative degree of the lag and the amount of release rate. Subject 1 shows the most separation between gestures (44.18 ms) and the highest release rate (.98), while the clusters of Subject 3 overlapped much more
(-6.63 ms) and are often unreleased (.54). However, all the subjects agree in the overall treatment of labial-coronal (/pt/, /pt'/, /p't/, /p't'/) vs. coronal-labial (/tp/, tp'/, /t'p/, /t'p'/) clusters. The latter always show more lag and higher release rate than the former (Lag: Subject 1: 55.85 vs. 32.5; Subject 2: 18.23 vs. 5.83; Subject 3: 7.58 vs. -20.8). The secondary articulation of C_1 or C_2 did not seem to play a major role in the relative timing of the stop gestures.[38]

5.5.2.2 A study of overlap in clusters

These timing asymmetries between clusters were further investigated in a study of gestural overlap in Russian stop clusters (Kochetov & Goldstein 2001). This study, using the data collected during the same EMMA sessions, extended the scope of the analysis to clusters with the dorsal /k/ as C_1 and C_2 (/kt/, /tk/, /kp/, /pk/), as well as to additional stimuli (with cross-word clusters varying in speech rate and with internal clusters in real words (Table 5.36)). Overall, about 250 tokens were collected per subject (10 clusters × 5 types × 5 repetitions) (not all conditions were examined for Subject 3).

Table 5.36

Boundaries	*Type of stimuli*	*Example*	*Gloss*
Cross-word	nonsense	e.g. tá[t#p]ápy	
	real, slow	e.g. brá[t#p]ádaja	'brother falling'
	real, fast	e.g. brá[t#p]ádaja	'brother falling'
Within-word	real, slow	e.g. o[tp]ál	'(it) fell off'

Below I report the results of our analysis that are relevant to this thesis and discuss some general implications.

The results of the study confirm that the order of place of articulation in a cluster is important for clusters with labials and coronals: coronal-labial clusters have more lag (less overlap) than labial-coronal sequences. This is an instance of the general effect of place order, found in other languages (Chitoran, Goldstein, & Byrd (in press) on Georgian; Surprenant & Goldstein 1998 on English). In these languages *front-to-back* clusters (/pt/, /pk/, /tk/) exhibit more overlap than *back-to-front* (/tp/, /kp/, /kt/) sequences. Here the term 'front-to-back' refers to clusters where the constriction for the first consonant is made in a more anterior position than for the second segment. For instance, in the cluster /pt/, the labial /p/ is articulated at the front, while the constriction of the second consonant, /t/, is further back. The term 'back-to-front' refers to the reverse order. For instance the cluster /tp/ is 'back-to-front', since the first consonant is articulated in a more posterior position in the oral cavity than /p/. This asymmetry is proposed to be perceptually motivated (Chitoran, Goldstein, & Byrd (in press)): even in the absence of the release burst of C_1 in front-to-back clusters can have an audible acoustic release (for example, due to the lip separation during the production of /p/). This is not true for C_1 in back-to-front

[38] However, the cluster /tp'/ for Subject 3 does not follow the pattern.

clusters, where release in an overlapped cluster is inaudible due to the constriction made at the front. This means that the recoverability of C_1 is at a greater risk in back-to-front clusters than in front-to-back sequences. Having a greater lag between the gestures in back-to-front clusters 'rescues' C_1 by providing more information about the segment.

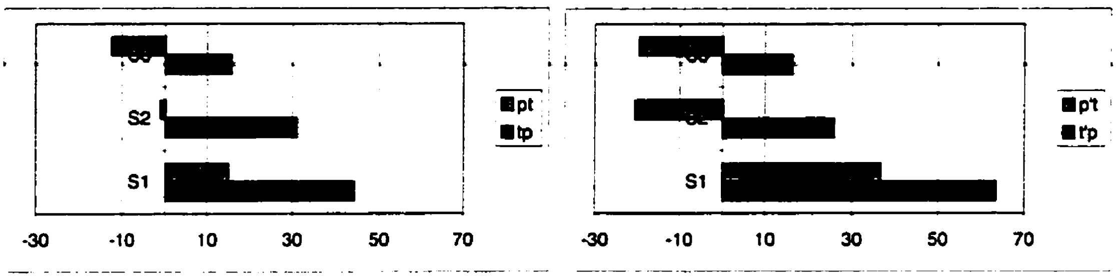
In this study it was found that while the place order effect was preserved in Russian clusters with labials and coronals, it was modified by another factor -- whether the cluster was contrastive in terms of palatalization with respect to its C_1 -- in sequences with dorsals (/pk/ vs. /kp/, /tk/ vs. /kt/). The study showed that the relative timing in Russian clusters did not depend on the presence or absence of the palatalized gesture. In other words, the degree of lag was statistically the same for both types of clusters of the same place order (CC and C'C).

The results for clusters with labials and coronals are presented in Table 5.37. The clusters are grouped in terms of their place order (front-to-back and back-to-front). For all the speakers the clusters labial + coronal show less lag (more overlap) than the clusters coronal + labial.

Table 5.37 The degree of lag in plain-plain and palatalized-plain hetero-organic clusters (additional conditions): Means and standard deviations

		Subject 1		Subject 2		Subject 3	
		Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*
front-to-back	pt	14.85	11.6	-1.08	14.84	-12.44	14.15
	p't	36.6	17.6	-20.4	6.49	-19.41	10
	mean	*25.73*		*-10.74*		*-15.93*	
back-to-front	tp	44.33	16.9	31	26.91	15.31	23.1
	t'p	63.6	28.8	25.73	37.26	16.31	8.96
	mean	*53.97*		*28.37*		*15.81*	
	overall mean	**39.85**		**8.81**		**-0.06**	

Figure 5.25 illustrates the differences between labial-coronal and coronal-labial clusters (grouped by secondary articulation of C_1). Despite the individual variation in duration values, all the subjects show the same relationship between front-to-back and back-to-front clusters.



a. b.

Figure 5.25 Degree of lag in plain-plain (a) and palatalized-plain (b) clusters with labials and coronals for three subjects

The factor of stimulus type (cross-word nonsense slow, real-slow, real-fast, and within-word real-slow) turned out to be significant as well. I examine the comparisons between some of the stimulus types here. For both subjects tested for speech rate (Subject 1 and Subject 2), the clusters in nonsense utterances showed significantly more lag than the clusters in real words pronounced at a faster rate (plain clusters Subject 1: nonsense 47.98, real fast 33.50; Subject 2 nonsense 42.47 real fast 17.19).[39] This is not surprising, since more co-articulation, and correspondingly more overlap, is usually found in faster and less formal speech (Browman & Goldstein 1989). These results also suggest a possible interpretation of the differences between the subjects. The higher lag found for Subject 1 in nonsense utterances is indicative of slower and less casual speech during the session.[40] More overlap characterizing the speech of Subject 3 could be due to the overall faster and more casual style.

In general, the findings show that there is more lag in coronal-labial clusters than in labial-coronal sequences, regardless of the secondary articulation of the consonants. In all clusters the degree of lag decreases in faster and more casual speech.

The results of this study, together with other works (Byrd 1992, Chitoran, Goldstein, & Byrd in press, Surprenant & Goldstein 1998, Zsiga 2000) provide evidence that gestural coordination in clusters is a very complex phenomenon influenced by a number of articulatory and perceptual

[39] Timing of gestures in nonsense utterances was not significantly different from that in within word clusters in real words.

[40] This style is also manifested in the lack of obligatory final devoicing of /b/ and /b'/ in some of the real utterances pronounced by Subject 1.

factors, and to a certain degree modified in a language-particular way (for instance, influenced by the language-specific system of contrasts and complexity in clusters). This will be further discussed in the final chapter.

5.5.2.3 Summary

Overall, for all speakers, the coronal-labial clusters show a higher degree of lag and higher release rate (and thus are more likely to be released) than the corresponding labial-coronal sequences.

5.5.2.4 Homorganic clusters

Recall that the acoustic study (Chapter 3) and our preliminary investigation of the articulatory results suggest that the first consonant in homorganic clusters is rarely released. Does this conclusion statistically hold true for all homorganic clusters of interest? Does it depend on the place and secondary articulation of the stops in the cluster? In this section I address these questions.

5.5.2.4.1 Overview

A homorganic cluster is articulatorily manifested in one steady constriction (or 'plateau') for the two consonants and no burst in the acoustic signal (cf. Figure 5.23a). A release of C_1 would involve an additional movement off the constriction and then back, i.e. two constrictions (or plateaus). The acoustic consequence of this articulatory manoeuvre is an acoustic burst between the two closures. An example of this is shown in Figure 5.26. The Lower Lip gesture of Subject 1 achieves its target (the constriction with the Upper Lip), remains there for less than 50 ms, and then is released. However, it is not returned to the original location, but moves toward the second constriction. As a result, we observe two constrictions, or two plateaus of the same gesture. Notice also the strong noise component, a release burst, in the signal between the two closures.

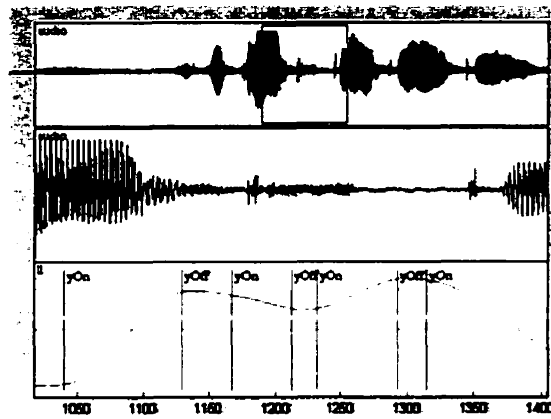


Figure 5.26 Gestures and acoustic signal during the production of the clusters /pp/ (in *tap papy*), Subject 1 (token n2_50)

5.5.2.4.2 Analysis

All the homorganic clusters (with labials, /pp/, /p'p/, and with coronals, /tt/, /t't/) in nonsense utterances were examined for the presence or absence of articulatory release of C_1 (total of 120 tokens: 8 utterances × 5 repetitions × 3 subjects). The labeling procedure (see section 5.2.3) allowed me to identify the number of constrictions in a given cluster. As with the hetero-organic clusters, the relative degree of articulatory release of a given cluster was measured. C_1 was considered released (assigned 1.00) if there were two constrictions (Lip Aperture or Tongue Tip). C_1 was treated as unreleased if it was determined that the gesture had one constriction (assigned .00). In determining this, reference was also made to the acoustic signal (burst). No duration measurements were taken.

The results for the three speakers were analyzed in an ANOVA with factors Place (labial and coronal), Pal (plain and palatalized) and a dependent variable Release (unreleased and released). The effect of subject was separately tested in an ANOVA with factor Speaker (Subject 1, Subject 2, and Subject 3), Pal, and the dependent variable Release.

Below I present the results. Means and standard deviations (averaged for 5 tokens) for lag in these clusters are given in Table 5.38. The results are organized by place of articulation of the stops in clusters and by speaker. .00 stands for no release of C_1 in the cluster and 1.00 means that the consonant was released 100% of the time.

Table 5.38 The articulatory release rate in homorganic clusters: Means

		Subject 1	Subject 2	Subject 3
labial	pp	0.20	0.00	0.00
	pp'	1.00	0.00	0.00
	p'p	1.00	0.00	0.00
	p'p'	1.00	0.00	0.30
	mean	*0.80*	*0.00*	*0.08*
coronal	tt	0.20	0.00	0.00
	tt'	0.20	0.20	0.00
	t't	0.80	0.00	0.00
	t't'	0.80	0.00	0.00
	mean	*0.50*	*0.05*	*0.00*
	overall mean	**0.65**	**0.03**	**0.04**

5.5.2.4.3 Results and discussion

The results of the statistical analysis show that neither Place [$F(1,24) = 482$, $p = .495$] nor Pal [$F(1,24) = 1.26$, $p = .275$] are significant, and there are no significant interactions of these factors. The main effect of Speaker is significant [$F(1,24) = 19.61$, $p = .000$]. No significant interactions with this factor are found. A Tukey HSD post-hoc reveals that there is a significant difference between Subject 1 and the other speakers (both $p = .000$). The values for Subject 2 and Subject 3 are not different ($p = .953$). This is not surprising, given the strikingly high release rate of Subject 1 compared to the virtual absence of release for the other two subjects.

The results reveal that the quality of consonants in the homorganic clusters (place and palatalization) does not influence the rate of release of C_1. Speakers treat homorganic clusters differently: while Subject 2 and Subject 3 almost never release the first segment, Subject 1 does it quite frequently. Recall that Subject 1 also showed a higher degree of lag in hetero-organic clusters (section 5.5.2.2). This degree of lag, however, is significantly lower under the fast speech condition. In order to determine whether the same holds for the homorganic cases, I examined the homorganic cross-word clusters /pp/, /tt/, and /t't/, as pronounced by Subject 1 in real words at a fast rate.[41]

The results for these three clusters in both slow and fast speech are shown in Table 5.39. While the drop in release rate in clusters /pp/ and /tt/ may not be significant (no statistical comparison was performed due to the small number of tokens), it is fairly substantial for /t'/ (a drop of 60%).

Table 5.39 The articulatory release rate in selected homorganic clusters at slow and fast rate: Means

	Slow	Fast
pp	.20	.00
tt	.20	.00
t't	.80	.20
overall mean	**.40**	**.07**

Figure 5.27 presents a released token of /pp/ (repeated from Figure 5.5) as pronounced by Subject 1 at a slow rate compared with a token of the same cluster pronounced fast. Notice the difference in the number of Lower Lip constrictions and the corresponding presence and absence of burst. It is also important to point to the overall duration differences: the distance between the first achievement of the target (yOff) and the last release (the fourth yOn) in the first token is almost 200 ms, while the duration of the constriction in the second token is slightly above 50 ms.

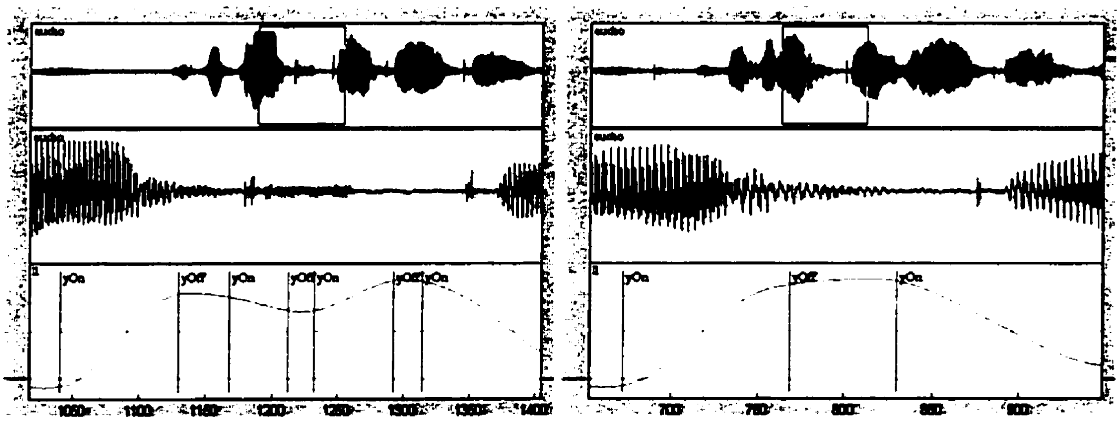


Figure 5.27 Gestures and acoustic signal during the production of the cluster /pp/ at a slow (in *tap papy*) and fast (in *grab padaja*) rate, Subject 1 (tokens n_20 and n4_13)

This suggests that the high release rate found in the slow nonsense homorganic tokens of Subject 1 can be attributed to the speech rate and style of this subject during the experiment. It is more common for speakers to leave the first stop in these clusters unreleased.

[41] Only utterances with these homorganic clusters were used in the fast condition.

5.5.2.5 C_2 in clusters

Most of our discussion has been devoted to the first consonant in clusters. This is not surprising, since the circumstances under which C_1 is released are very complex. Unlike C_1, the second stop of a cluster, C_2, is followed by a vowel, and thus is always released. The presence of burst for C_2 under any circumstances is likely to make this consonant substantially easier to recover compared to the first segment.

5.5.2.6 Summary: Clusters

The comparison of hetero-organic and homorganic clusters with plain and palatalized labials and coronals shows that they differ in the degree to which the first stop in the cluster can be released both articulatorily and acoustically. C_1 is found to be released more often before a consonant of a different place of articulation than before a homorganic segment. In addition, hetero-organic clusters vary in the degree of lag between the two primary gestures. More lag is observed in back-to-front (coronal + labial) sequences rather than in front-to-back (labial + coronal) clusters. In addition, in faster and more casual speech the first gesture tends to overlap more with the following gesture. This leads to less frequent release bursts for the first stop.

The striking asymmetry between hetero- and homorganic clusters confirms the results of previous acoustic studies of Russian stop clusters (Chapter 3; Zsiga 2000).

5.5.3 Summary and predictions

In this section I summarize the overall results for articulatory release and acoustic burst in single stops and in clusters. I also make predictions about the ease of perceptual recoverability of stops in different positions based on the burst properties discussed here. These predictions are examined in detail in the study reported in Chapter 6.

We have seen that there are substantial differences between consonants and environments with respect to the quality of release and burst, and also with respect to its presence or absence (section 5.5). The burst quality differences point to asymmetries among the four consonants. In

(5.4) I state these asymmetries as relative recoverability scales that hold that, based on their acoustic burst properties, coronals should be easier to recover than corresponding labials, and palatalized segments are likely to be better perceived than their plain counterparts.

(5.4) Acoustic burst quality

a. /t'/, /t/ > /p/, /p'/

b. /t'/ > /t/

c. /p'/ > /p/

This perceptual recoverability scale should hold for a stop in any environment (single, initial or final, C_1 or C_2 in a cluster). Other asymmetries refer to particular positions in a cluster. We saw that while the second consonant in a cluster is always audibly released, this is not always true for the first segment (section 5.5.2). This is stated in (5.5): the quality of C_2 should be easier to recover than that of C_1.

(5.5) Presence/absence of audible release depending on position in cluster

$C_2 > C_1$, or C/C_V > C/V_C

We also saw differences in the release rate of C_1 depending on whether the cluster is hetero-organic (section 5.5.2.2) or homorganic (section 5.5.2.4). The hetero-organicity condition is given in (5.6). It states that a stop followed by a consonant of a different place of articulation should be easier to recover than when it is before a homorganic segment.

(5.6) Presence/absence of audible release depending on the place of the following consonant

$C_i/\_C_j > C_i/\_C_i$

Finally, the results reveal a place order effect with coronal-labial clusters having a higher lag than labial-coronal sequences (section 5.5.2.2). Since the burst contains information about both place and the secondary articulation of C_1, we may expect that the place and palatalization of the first consonant in labial-coronal clusters (/pt/, /pt'/, /p't/, /p't'/) will be harder to recover perceptually than the quality of C_1 in coronal-labial stop sequences (/tp/, /tp'/, /t'p/, /t'p'/) (5.7).

(5.7) Degree of lag depending on the place of the following consonant

t, t'/_p, p' > p, p'/_t, t'

It is important to mention that the first condition, the quality of the acoustic burst, can be applied to any of the other factors. Thus, under the same positive degree of lag in clusters /pt/, /p't/, /tp/, and /t'p/, we can expect that the palatalized /p'/ or /t'/ as C_1 should be more easily recovered than their plain counterparts, /p/ and /t/; also, coronals /t/ and /t'/ are more likely to be correctly identified than the corresponding labials, /p/ and /p'/.

In sum, the differences found in this section have important consequences for the perception of the plain and palatalized labial stops.

The predictions discussed above are based on release and burst properties only. Other factors, such as primary and secondary gesture differences and the corresponding acoustic transition may modify or even override these factors in perception. These predictions and the relative importance of various articulatory and acoustic factors will be further tested perceptually in Chapters 6 and 7.

All articulatory results for Subject 3, summarized below (Tables 5.40 and 5.41), will be used in correlations between articulation and perception in Chapter 7.

Table 5.40. Results of articulatory experiments for Subject 3: Tongue Body magnitude and slope (Tby and TBx) at various points in time

a. Single consonants

Stimuli	TBy magnitude				TBx magnitude				TBy slope		TBx slope	
	VC1	VC2	CV1	CV2	VC1	VC2	CV1	CV2	VC	CV	VC	CV
ta papy	-4.3	-4.7	-6.3	-6.1	35.7	39.3	41.3	40.4	-0.3	0.22	-3.6	0.92
ta p'apy	-3.9	-1.6	1.64	-1.8	35.2	35.2	29.6	32.3	2.32	-3.4	-0	-2.7
ta tapy	-4.4	-5.1	-5.6	-4.9	36.7	35.8	36.2	38.4	-0.7	0.64	0.86	-2.3
ta t'apy	-2.1	2.52	2.78	-0.9	35.6	33.6	30.4	34.2	4.62	-3.7	2.02	-3.8
tap apy	-3.7	-3.3	-4.4	-4.5	36.2	38.8	39.9	39.2	0.43	-0.1	-2.7	0.72
tap' apy	-1.2	0.98	-2.9	-4	33.8	33.4	37.9	39	2.2	-1.2	0.4	-1
tat apy	-5.0	-4.7	-3.8	-4	35.1	33.8	34.7	36.4	0.24	-0.2	1.3	-1.7
tat' apy	-1.4	2.48	0.54	-2	35.3	32.6	32.8	36.7	3.86	-2.5	2.7	-3.9

b. Onset and coda consonants in clusters

Stimuli	Onset						Coda					
	TBy magnitude		TBx magnitude		TBy slope	TBx slope	TBY magnitude		TBx magnitude		TBy slope	TBx slope
	CV1	CV2	CV1	CV2	CV	CV	VC1	VC2	VC1	VC2	VC	VC
tap papy	-6.3	-6.1	41.3	40.4	0.22	0.92	-3.6	-3.4	36.1	38.5	0.24	-2.4
tap' papy	-6.3	-6.1	41.3	40.4	0.22	0.92	-2.1	0	35.3	34.9	2.08	0.4
tap p'apy	1.64	-1.8	29.6	32.3	-3.4	-2.7	-2.7	-0.3	34.5	33.8	2.43	0.65
tap' p'apy	1.64	-1.8	29.6	32.3	-3.4	-2.7	0.3	2.77	33.9	32.1	2.47	1.77
tap tapy	-5.6	-4.9	36.2	38.4	0.64	-2.3	-4.4	-5.1	36.5	36.6	-0.7	-0.1
tap' tapy	-5.6	-4.9	36.2	38.4	0.64	-2.3	-1.8	0.32	34.4	33.1	2.16	1.26
tap t'apy	2.78	-0.9	30.4	34.2	-3.7	-3.8	-3.1	-0.5	35	35.4	2.62	-0.4
tap' t'apy	2.78	-0.9	30.4	34.2	-3.7	-3.8	-1.9	0.62	34.4	33.1	2.56	1.34
tat papy	-6.3	-6.1	41.3	40.4	0.22	0.92	-3.7	-3.3	34.4	34	0.44	0.42
tat' papy	-6.3	-6.1	41.3	40.4	0.22	0.92	-0.6	3.58	34.5	31.6	4.16	2.96
tat p'apy	1.64	-1.8	29.6	32.3	-3.4	-2.7	-3.6	-1.8	34.9	32.5	1.78	2.45
tat' p'apy	1.64	-1.8	29.6	32.3	-3.4	-2.7	-0.2	4.36	35	31.8	4.58	3.18
tat tapy	-5.6	-4.9	36.2	38.4	0.64	-2.3	-4.3	-3.2	35.4	33.3	1.12	2.08
tat' tapy	-5.6	-4.9	36.2	38.4	0.64	-2.3	-1.3	1.93	34.3	31.7	3.23	2.6
tat t'apy	2.78	-0.9	30.4	34.2	-3.7	-3.8	-2.4	0.9	35.1	32.7	3.28	2.42
tat' t'apy	2.78	-0.9	30.4	34.2	-3.7	-3.8	-0.6	3.46	35.9	32.8	4.04	3.1

Table 5.41. Results of articulatory experiments for Subject 3: Release rate, burst quality, lag, Tongue Body differences in magnitude, Lip Aperture and Tongue Tip magnitude

a. Single consonants

Stimuli	Relea-se rate	Burst Qual./ Relea-se rate	TBy differ-ence	TBx differ-ence	TBy difference		TBx difference		Difference		
			total	total	C vs. C'	lab vs. cor	C vs. C'	lab vs. cor	LA	TT	TT
ta papy	1	1	86.6	97.8	15.8	2.46	24.4	12.1	19.3		
ta p'apy	1	2	86.6	97.8	15.8	7.94	24.4	10.2	18.2		
ta tapy	1	2	86.6	97.8	22.3	2.46	13.3	12.1		-4.4	13.3
ta t'apy	1	3	86.6	97.8	22.3	7.94	13.3	10.2		-1.1	11.1
tap apy	1	1	63	61.4	8.82	3.8	9.83	14	16.8		
tap' apy	1	2	63	61.4	8.82	7.1	9.83	9.78	17.1		
tat apy	1	2	63	61.4	17.2	3.8	3.58	14		-4.3	12.2
tat' apy	1	3	63	61.4	17.2	7.1	3.58	9.78		-1.2	13.1

b. Onset consonants in clusters

Stimuli	Relea-se rate	Burst Qual./ Relea-se rate	TBy differ-ence	TBx differ-ence	TBy difference		TBx difference	
			total	total	C vs. C'	lab vs. cor	C vs. C'	lab vs. cor
tap papy	1	1	52.6	74	12.3	1.94	19.7	7.1
tap' papy	1	1	52.6	74	12.3	1.94	19.7	7.1
tap p'apy	1	2	52.6	74	12.3	1.28	19.7	7.46
tap' p'apy	1	2	52.6	74	12.3	1.28	19.7	7.46
tap tapy	1	2	52.6	74	12.4	1.94	9.98	7.1
tap' tapy	1	2	52.6	74	12.4	1.94	9.98	7.1
tap t'apy	1	3	52.6	74	12.4	1.28	9.98	7.46
tap' t'apy	1	3	52.6	74	12.4	1.28	9.98	7.46
tat papy	1	1	52.6	74	12.3	1.94	19.7	7.1
tat' papy	1	1	52.6	74	12.3	1.94	19.7	7.1
tat p'apy	1	2	52.6	74	12.3	1.28	19.7	7.46
tat' p'apy	1	2	52.6	74	12.3	1.28	19.7	7.46
tat tapy	1	2	52.6	74	12.4	1.94	9.98	7.1
tat' tapy	1	2	52.6	74	12.4	1.94	9.98	7.1
tat t'apy	1	3	52.6	74	12.4	1.28	9.98	7.46
tat' t'apy	1	3	52.6	74	12.4	1.28	9.98	7.46

c. Coda consonants in clusters

Stimuli	Release rate	Burst Qual./ Release rate	Lag	TBy difference	TBx difference	TBy difference		TBx difference	
				total	total	C vs. C'	lab vs. cor	C vs. C'	lab vs. cor
tap papy	0	0		5.85	4.55	4.97	0.2	4.46	6.14
tap' papy	0	0		5.85	4.55	4.97	5.08	4.46	4.02
tap p'apy	0.2	0.2		5.94	1.67	6.04	2.4	2.22	1.8
tap' p'apy	0.33	0.66		5.94	1.67	6.04	2.11	2.22	1.41
tap tapy	0.2	0.2	-15	6.02	3.73	7.9	1.94	5.6	4.42
tap' tapy	0.2	0.4	-17	6.02	3.73	7.9	2.15	5.6	1.47
tap t'apy	0	0	-19	3.32	2.15	2.06	2.1	2.92	2.8
tap' t'apy	0.3	0.6	-32	3.32	2.15	2.06	4.2	2.92	1.76
tat papy	0.5	1	5	5.85	4.55	9.96	0.2	2.54	6.14
tat' papy	1	3	16	5.85	4.55	9.96	5.08	2.54	4.02
tat p'apy	0.25	0.5	-12	5.94	1.67	9.5	2.4	0.73	1.8
tat' p'apy	1	3	20.8	5.94	1.67	9.5	2.11	0.73	1.41
tat tapy	0	0		6.02	3.73	8.11	1.94	2.65	4.42
tat' tapy	0	0		6.02	3.73	8.11	2.15	2.65	1.47
tat t'apy	0	0		3.32	2.15	4.36	2.1	0.88	2.8
tat' t'apy	0	0		3.32	2.15	4.36	4.2	0.88	1.76

Chapter 6. A perceptual study

6.0 Introduction

Based on the results of the experiments reported in Chapter 5 (and Chapters 3 and 4) a number of predictions have been made about the perception of the plain-palatalized contrast. Recall from Chapter 1 that I am assuming a model where perception is viewed as a process of extracting information about articulatory gestures/features from an acoustic signal. This implies that a listener would be less successful at this task when the information about the gestural/featural structures is minimal or inconclusive. This allows me to make a number of specific predictions as to when listeners are more or less likely to fail at recovering the underlying gestures/features.

Based on the articulatory results and this hypothesis I predict the following. We should expect significant differences in the perception of palatalization and according to environment. We are likely to find better identification of the plain-palatalized contrast in the onset environment than in the coda environment. There should be no major difference between the perception of onset single consonants and onset consonants in clusters. However, we would expect to find that the contrast is perceived less reliably in the preconsonantal rather than final position. In the preconsonantal environment, there should be more confusion before homorganic consonants than before consonants of a different place of articulation (mainly for coronals). Also, there should be less confidence in the identification of a consonant followed by a palatalized segment than of one followed by a plain segment. The worst environment for the correct identification of a palatalized consonant should be before a homorganic plain consonant. The worst position for the recognition of a plain consonant should be before a homorganic palatalized consonant.

We should also expect differences with respect to the place of articulation of the palatalized consonant. The palatalized labial /p'/ should be more likely to be confused with the plain /p/ than /t'/ with /t/. This should be particularly true of the coda environment.

The goal of this chapter is to test the predictions summarized above through a series of perceptual experiments. Since our hypothesis is based on the assumption of a general or 'universal' ability of listeners to recover gestures/features, it is desirable to control for language-particular effects on perception. To do so I test both Russian and Japanese listeners. It is expected that they would perform similarly in terms of overall recoverability of gestures/features, but listeners of different languages may be influenced by language-particular factors such as the overall system of phonological contrasts, phonotactic knowledge, and familiarity with the presented gestures. It is the similarity between the two groups that of primarily interest in this thesis. The two groups, Russian listeners and Japanese listeners, performed a forced choice phoneme identification task. In addition, they were asked to write down the stimuli. This was done to examine other strategies listeners might exploit when representing the recovered gestures/features. In addition, I tested perception of Russian listeners in noise to investigate what gestural and acoustic properties are most and least robust, and what environments are most affected.

We find that the results generally confirm our predictions. Both groups of listeners under various conditions show a high rate of correct identification of stops in syllable onset position and much poorer recognition in the preconsonantal environment. The context before palatalized segments is the most error-prone. The writing task reveals that listeners are less successful in recovering the timing of the palatalized gesture/feature with respect to other gestures/features.

The results from both Russian and Japanese subjects provide additional evidence supporting the earlier findings, but they also show that palatalization can be interpreted by listeners in a variety of ways. When presented with a stimulus with a palatalized consonant, a listener may correctly segment the input as a palatalized consonant or s/he may attribute palatalization to a neighbouring (mostly preceding) consonant or vowel. A consonant in the coda environment is particularly hard to identify in terms of palatalization (but also in terms of other features). It is more often confused with a following segment, when present. At the same time, its palatalized quality may be attributed to the preceding vowel and consonant.

The chapter is organized as follows. Sections 6.1 and 6.2 examine the experiments on phoneme identification by Russian listeners given clean (without noise) and noisy stimuli respectively. The identification by Japanese subjects is presented in section 6.3. Sections 6.3 and 6.4 are devoted to the writing task by Russian and Japanese listeners. Overall results are summarized in section 6.5.

6.1 Identification task (clean stimuli): Russian listeners

6.1.0 Introduction

In this section I investigate the perception of plain and palatalized stops by Russian listeners without noise. The task is phoneme identification.

6.1.1 Experimental setup, materials and procedure

Twenty native speakers of Russian[42] were involved in this experiment. None of the listeners had any known hearing disorders. All subjects were tested individually. Test utterances consisted of nonsense phrases of the type *taC₁# C₂apy* from the articulatory experiment as pronounced by Speaker 1 (repeated in Table 6.1). They included four consonants (C_1 and C_2: /p/, /p'/ /t/, and /t'/) in 5 onset (prevocalic) and 5 coda (postvocalic) environments. All clusters are across word boundaries, since not all combinations are possible within words. I refer to the environment *a_#a* as the final coda, even though the word-final consonant is followed by a vowel. As we will see, the results for this environment are not very different from other coda positions.[43] The first two

[42] The subjects were 11 males and 9 females.

[43] It appears that at least some of the consonants were not re-syllabified or were ambi-syllabic in this position. The details of word boundary syllabification are beyond the scope of this study. See Bondarko 1977 for a discussion of syllable structure in Russian in general (122-151) and of syllabification at word boundaries (137-139).

vowels were both stressed, but stronger stress tended to be on the second vowel. In all of the cases (except a#_a, where /ta/ was a part of the carrier phrase and was reduced to a schwa), however, the quality of the less stressed [a] was different from that of a typical reduced vowel.

Table 6.1 Stimuli for the perception experiment

Environment		p	p'	t	t'
Onset (#/C_V)	a#_a	ta papy	ta p'apy	ta tapy	ta t'apy
	t#_a	tat papy	tat p'apy	tat tapy	tat t'apy
	p#_a	tap papy	tap p'apy	tap tapy	tap t'apy
	t'#_a	tat' papy	tat' p'apy	tat' tapy	tat' t'apy
	p'#_a	tap' papy	tap' p'apy	tap' tapy	tap' t'apy
Coda (V_#/C)	a_#a	tap apy	tap' apy	tat apy	tat' apy
	a_#t	tap tapy	tap' tapy	tat tapy	tat' tapy
	a_#p	tap papy	tap' papy	tat papy	tat' papy
	a_#t'	tap t'apy	tap' t'apy	tat t'apy	tat' t'apy
	a_#p'	tap p'apy	tap' p'apy	tat p'apy	tat' p'apy

The audio signals were extracted from MATLAB files and was saved as audio files (.wav). The utterances were excised from the carrier phrases using SoundEdit software. Two sets of stimuli were prepared: without noise ('clean' stimuli) and with noise ('noisy' stimuli). For the latter stimuli I added white noise using SoundEdit (option 'Add Noise'). The onset and offset of noise were tapered to avoid an audible click. Each token had a duration of about 600 ms. The level of noise, determined in a pilot study with three subjects, was considered to be adequate to induce a large number of errors, while maintaining identification above the chance level. The same stimuli were used for onset and coda consonants in clusters. Each utterance was repeated five times. The repetitions corresponded to five different tokens of the same utterance. The stimuli were organized into four 100-trial blocks (clean coda utterances, noisy coda utterances, clean onset utterances, and noisy onset utterances) and presented using the program PsyScope (Cohen, MacWhinney, Flatt, and Provost 1993) in random order with an inter-stimulus interval of 1500 ms. Listeners were prompted visually with an asterisk on the computer screen (for 500 ms). This was followed by an audio signal. Four keys were labeled in Cyrillic to identify the phonemes /p/, /p'/, /t/, and /t'/ (as *pa, pja, ta, tja* in onset and as *p, p', t, t'* in coda). Listeners were instructed to press the key corresponding to the phoneme they heard as soon as possible.

They were told to make a choice even if they were in doubt. If no key was pressed within 4000 ms, the next stimulus was presented. There were 10-second breaks between the blocks. The program kept track of the identification responses and response times (RT). Before each session the subjects were given written examples of possible stimuli and underwent a practice block (about 50 tokens).[44]

6.1.2 Analysis

For each subject I calculated mean identification rates (proportion correct out of 1.00) based on averages of 5 tokens per utterance. This yielded a total of 8000 identification tokens (40 utterances × 5 repetitions × 2 conditions × 20 listeners). There were 49, or .63% (31 clean and 18 noisy) cases when subjects did not respond to a stimulus. The same procedure was applied to the RT data. In the analysis of RT only positive responses (i.e. all but errors in identification) were considered. Due to the high error rate in the noisy data, only clean tokens were examined for RT. To normalize the data I excluded all values lower or higher than two standard deviations for each subject (based on the overall mean RT values per subject). A total of 187 clean tokens (about 9 tokens per subject) were excluded.

The results were statistically analyzed using the *Statistica* package. Clean and noisy stimuli were examined separately. Within each group I ran separate repeated-measures ANOVAs for three sub-groups: single consonants (onset and coda together), onset consonants in clusters and coda consonants in clusters. This division of the data allowed me to examine the full interaction of factors. The within-subject factors used in the ANOVAs for single consonants were Place (2 levels: labial vs. coronal), Palatalization (*Pal,* 2 levels: plain vs. palatalized), and Environment (*Env*, 2 levels: onset vs. coda). The factors used for consonants in clusters (onset and coda) were Place (2 levels: labial vs. coronal), Palatalization (*Pal*, 2 levels: plain vs. palatalized), Environment Place (*EnvPlace* (2 levels: labial vs. coronal), and Environment Palatalization

[44] In addition to that, prior to this experiment the subjects performed a task of writing down the same stimuli. This

(*EnvPal*, 2 levels: plain vs. palatalized). In both analyses the dependent variable corresponded to 20 repeated measurements of identification per subject. I used a Tukey HSD post-hoc test to test difference among means in significant interactions. In order to compare the results for single consonants and consonants in clusters, the total mean values for each utterance were analyzed in multivariate ANOVAs for onset and coda environments. The factors were Type (2 levels: single vs. cluster), Place (2 levels: labial vs. coronal), and Pal (2 levels: plain vs. palatalized). The same analysis was performed for both correct identification and RT.

In the following section I examine and discuss the results of the first part of the experiment: perception of stimuli without noise. To put our hypothesis in terms of statistical results, we should find the following. For single consonants the most likely factor to be significant is Env (onset vs. coda). If the effect is limited to the palatalized consonants, we expect to find a significant interaction of Env × Pal. Further, if the effect is mostly limited to /p'/ and is strong enough, we may encounter a significant three-way interaction Env × Place × Pal, given our articulatory findings. The interaction of Place × Pal should also be significant if each consonant shows a different identification rate (and RT) irrespective of the environment.

For onset consonants in clusters we expect no major effect of any of the factors (it should be noted that in the articulatory study I did not examine these consonants in detail).

For coda consonants in clusters we are likely to find a number of significant effects (especially EnvPal) and interactions (especially Pal × EnvPal). The asymmetries between the palatalized consonants and place of articulation of the following consonant are also likely to play some role giving rise to additional interactions of factors.

6.1.3 Overall results

In this section I give a general overview of the results. The results by major groups of stimuli as well as discussion of these results are presented in more detail in later sections. Means and

---

study is reported as a separate experiment in section 6.4.

standard deviations for correct identification of the four stops, based on the average for all subjects, are presented in Tables 6.2-6.4. The stimuli are divided into groups by syllable position. The consonants in question are given in bold. A score of .99 for the utterance *ta papy* means that the consonant /p/ in this utterance was identified as /p/ in 99% of the cases (based on the results from 20 subjects, 5 tokens per subject). The chance level for each environment was .25.

Table 6.2 Correct identification: Single consonants (means and standard deviations)

initial	*Mean*	*SD*	*final*	*Mean*	*SD*
ta papy	0.99	0.04	tap apy	0.80	0.25
ta p'apy	0.98	0.06	tap' apy	0.38	0.28
ta tapy	0.98	0.06	tat apy	0.97	0.07
ta t'apy	1	0.00	tat' apy	0.83	0.19
	.99	*0.04*		*0.75*	*0.20*

Table 6.3 Correct identification: Consonants in clusters (means and standard deviations)

onset	*Mean*	*SD*	*coda*	*Mean*	*SD*
tat papy	0.9	0.15	tap tapy	0.9	0.18
tap papy	0.98	0.06	tap papy	0.75	0.26
tat' papy	0.92	0.12	tap t'apy	0.68	0.30
tap' papy	0.95	0.11	tap p'apy	0.35	0.28
tat p'apy	1	0.00	tap' tapy	0.42	0.32
tap p'apy	0.98	0.06	tap' papy	0.34	0.27
tat' p'apy	0.91	0.14	tap' t'apy	0.37	0.28
tap' p'apy	0.98	0.06	tap' p'apy	0.72	0.30
tap tapy	0.98	0.06	tat papy	0.52	0.20
tat tapy	1	0.00	tat tapy	0.72	0.19
tap' tapy	0.82	0.21	tat p'apy	0.41	0.22
tat' tapy	0.88	0.14	tat t'apy	0.16	0.17
tap t'apy	1	0.00	tat' papy	0.88	0.24
tat t'apy	1	0.00	tat' tapy	0.53	0.33
tap' t'apy	0.99	0.04	tat' p'apy	0.86	0.22
tat' t'apy	1	0.00	tat' t'apy	0.86	0.27
	0.96	*0.07*		*0.62*	*0.24*

The total average identification rate was .79 (.15)[45]. The first number here stands for the mean value and the number in parentheses is the standard deviation. The details of inter-listener variation were not examined in this study. The rate for single consonants was .87 (.12) and for consonants in clusters was .79 (.16). Within the single consonants a rate of .99 (.04) was found for the onset (initial) and a rate of .75 (.20) for the coda (final). In clusters the onset consonants had a rate of .96 (.01) and the coda segments a rate of .62 (.24). Collapsing all the groups into onset and coda classes, the total rate for the onset is .98 (.06), and the total rate for the coda is .69 (.22).

Tables 6.4-6.7 present the identification matrices for each of the four consonants, i.e. how many token of a given consonant in a given environment were perceived as /p/, /p'/, /t/, and /t'/. The identification by consonant is combined with the identification by place (labial or coronal) and palatalization (plain or palatalized). The maximum number per environment are given in bold. Thus, from Table 6.4 we can see that /p/ in the final coda environment (a_#a) was perceived as /p/ 81 times (out of 99 responses; 1 missing); in all other cases it was identified as /p'/ (6 times), as /t/ (10 times), and as /t'/ (twice). Describing the same in terms of place and palatalization, we can say that in 87 cases /p/ was identified as labial (and in 12 cases as coronal) and in 91 cases it was perceived as plain (vs. 9 identifications as palatalized). Below I present the total and percentage for each percept. For instance, in total /p/ was perceived as /p/ in 82% of the cases; of all other consonants, it was mostly confused with /p'/ (10%), etc. The patterns of identification and confusion will be examined further in discussion sections.

Table 6.4 Confusion matrix: /p/ (tokens)

		Consonant				Place		Palatalization	
		p	p'	t	t'	lab	cor	npal	pal
single	a#_a	**98**	1	0	0	**99**	0	**98**	1
	a_#a	**81**	6	10	2	**87**	12	**91**	8
CC-onset	t_a	**88**	1	8	1	**89**	9	**96**	2
	p_a	**96**	0	1	1	**96**	2	**97**	1
	t'_a	**92**	3	3	1	**95**	4	**95**	4
	p'_a	**94**	5	0	1	**99**	1	**94**	6
CC-coda	a_t	**75**	12	9	4	**87**	13	**84**	16
	a_p	**90**	5	5	0	**95**	5	**95**	5
	a_p'	35	**52**	8	5	**87**	13	43	**57**
	a_t'	**67**	19	5	8	**86**	13	**72**	27
	total	***816***	*104*	*49*	*23*	***920***	*72*	***865***	*127*
	ratio	***0.82***	*0.10*	*0.05*	*0.02*	***0.93***	*0.07*	***0.87***	*0.13*

[45] with the lowest result .68 by subject BT and the highest .88 by the subject ML.

Table 6.5 Confusion matrix: /p'/ (tokens)

		Consonant				Place		Palatalization	
		p	p'	t	t'	lab	cor	npal	pal
single	a#_a	0	92	0	2	92	2	0	94
	a_#a	32	36	10	20	68	30	42	56
CC-onset	t'_a	2	88	1	6	90	7	3	94
	p'_a	0	97	0	2	97	2	0	99
	t_a	0	99	0	0	99	0	0	99
	p_a	1	99	0	0	100	0	1	99
CC-coda	a_t	49	34	6	11	83	17	55	45
	a_p	42	42	7	9	84	16	49	51
	a_p'	19	72	2	7	91	9	21	79
	a_t'	41	38	4	17	79	21	45	55
	total	186	697	30	74	883	104	216	771
	ratio	0.19	0.71	0.03	0.07	0.89	0.11	0.22	0.78

Table 6.6 Confusion matrix: /t/ (tokens)

		Consonant				Place		Palatalization	
		p	p'	t	t'	lab	cor	npal	pal
single	a#_a	0	0	96	2	0	98	96	2
	a_#a	2	0	95	1	2	96	97	1
CC-onset	p_a	3	0	97	0	3	97	100	0
	t_a	0	0	99	0	0	99	99	0
	p'_a	3	0	79	14	3	93	82	14
	t'_a	5	0	87	8	5	95	92	8
CC-coda	a_p	15	3	52	29	18	81	67	32
	a_t	16	2	72	10	18	82	88	12
	a_p'	10	21	41	28	31	69	51	49
	a_t'	10	7	16	67	17	83	26	74
	total	64	33	734	159	97	893	798	192
	ratio	0.06	0.03	0.74	0.16	0.10	0.90	0.81	0.19

Table 6.7 Confusion matrix: /t'/ (tokens)

		Consonant				Place		Palatalization	
		p	p'	t	t'	lab	cor	npal	pal
single	a#_a	0	0	1	99	0	100	1	99
	a_#a	1	1	15	83	2	98	16	84
CC-onset	p'_a	0	1	0	99	1	99	0	100
	t'_a	0	0	0	100	0	100	0	100
	p_a	0	0	0	100	0	100	0	100
	t_a	0	0	0	100	0	100	0	100
CC-coda	a_p	5	1	6	88	6	94	11	89
	a_t	13	15	20	52	28	72	33	67
	a_p'	2	8	4	86	10	90	6	94
	a_t'	5	2	7	86	7	93	12	88
	total	26	28	53	893	54	946	79	921
	ratio	0.03	0.03	0.05	0.89	0.05	0.95	0.08	0.92

Further, in Tables 6.8-6.9 I present the results for RT by the three groups of stimuli. The total average RT was 1464 (237) ms. The mean result for single consonants was 1449 (320) ms and for consonants in clusters was 1474 (328). Within the single consonants a rate of 1406 (279) ms was determined for the onset (initial) and a rate of 1492 (363) ms for the coda (final). In clusters the onset consonants had a rate of 1489 (298) and the coda segments a rate of 1459 (358) ms. Collapsing all the groups into onset and coda classes, the rate total for the onset is 1448 (288) ms, and the total rate for the coda is 1476 (360) ms.

Table 6.8 Response Time: Single consonants (means and standard deviations)

initial	Mean	SD	final	Mean	SD
ta papy	1433	299	tap apy	1509	344
ta p'apy	1357	268	tap' apy	1621	493
ta tapy	1455	257	tat apy	1406	332
ta t'apy	1380	287	tat' apy	1433	277
	1406	*278*		*1492*	*362*

Table 6.9 Response Time: Consonants in clusters (means and standard deviations)

onset	Mean	SD	coda	Mean	SD
tat papy	1452	262	tap tapy	1462	298
tap papy	1391	275	tap papy	1341	326
tat' papy	1455	245	tap t'apy	1575	334
tap' papy	1504	376	tap p'apy	1547	579
tat p'apy	1430	247	tap' tapy	1638	402
tap p'apy	1332	224	tap' papy	1453	376
tat' p'apy	1447	315	tap' t'apy	1597	530
tap' p'apy	1381	302	tap' p'apy	1307	288
tap tapy	1562	363	tat papy	1388	322
tat tapy	1477	260	tat tapy	1369	313
tap' tapy	1747	332	tat p'apy	1478	389
tat' tapy	1705	302	tat t'apy	1587	350
tap t'apy	1404	248	tat' papy	1380	332
tat t'apy	1488	318	tat' tapy	1474	286
tap' t'apy	1612	358	tat' p'apy	1412	335
tat' t'apy	1438	329	tat' t'apy	1340	270
	1489	*297*		*1459*	*358*

In the next sections I present the results in more detail.

6.1.4 Single consonants

6.1.4.1 Results

Table 6.10 shows the ANOVA results for the identification of onset consonants in clusters. Figure 6.1 presents the mean values for each of the onset consonants in various clusters.

Table 6.10 Correct identification: Single consonants (statistics)

	F	p
Place	64.4	0.000
Pal	30.65	0.000
Env	87.16	0.000
Place × Pal	9.04	0.007
Place × Env	59.97	0.000
Pal × Env	32.64	0.000
Place × Pal × Env	5.22	0.034

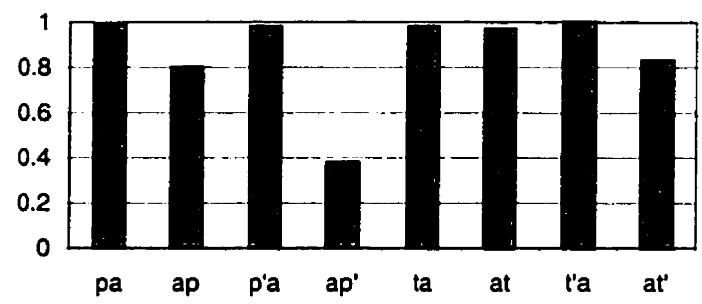


Figure 6.1 Correct identification: Single consonants

All of the factors, Place, Pal, and Env, are significant. As we expected, Env is the most important factor: it has the highest F value of all the factors and interactions. Correct identification of consonants in coda (.75) is significantly lower than in onset (.99) (Env: [$p < .001$]). Palatalized consonants are less correctly identified than their plain counterparts (.80 vs. .94) (Pal: [$p < .001$]).[46] The third factor is Place: labials were less reliably identified than coronals (.79 vs. .95).[47]

All two- and three-way interactions of the factors are also significant. The interaction Place × Env is highly significant. This interaction means that perception of single consonants is sensitive to the place of articulation in a given environment. A Tukey HSD post-hoc test determined that initial labials are identified significantly better than final labials (.99 vs. .59) ($p = .000$). The same was determined for initial and final coronals ($p = .000$).

[46] This effect, however, was influenced by the lower identification of the palatalized segments in the coda. Note that there was little difference between plain and palatalized consonants in the onset position.
[47] This holds true for the final consonants only.

The interaction Place × Pal shows the sensitivity to the particular consonant rather than parameters of place or palatalization, and thus can be referred to as the Consonant factor. This effect is due to the significant difference between /p'/ and all other consonants (p = .000). Note that /p'/ has the lowest recognition rate: .68.

The interaction Pal × Env reflects the dependence of the responses on the secondary articulation in a given environment. While final plain and palatalized consonants were significantly different from each other (.89 vs. .61) (p = .000), the same consonants in the initial position showed no difference (both .99).

The three-way interaction (Place × Pal × Env) is also significant. This indicates that the consonants in question are different depending on the environment they are in. The post-hoc test, however, determined that the effect comes mainly from /p'/, which is significantly different from all other consonants (p = .000) (See Figure 6.1). In addition to this, initial and final variants of /p/ were different from each other (p = .042).

In sum, all the factors (Place, Pal, and Env) and all interactions of these factors are significant. Most of the interactions are due to the lower recognition of /p'/ in the coda, as well as a relatively low rate for /p/ and /t'/ in this environment.

Table 6.11 shows the ANOVA results of the reaction time (RT) for onset consonants in clusters. Here and further I present only significant factor interactions. Figure 6.2 provides the mean values for each of the onset consonants in various clusters.

Table 6.11 RT: Single consonants (statistics)

	F	p
Place	**5.24**	**0.035**
Pal	.46	0.507
Env	**4.68**	**0.045**
Place × Env	**5.22**	**0.035**

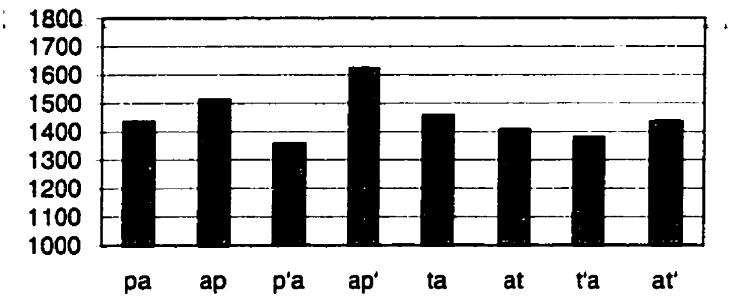


Figure 6.2 RT (ms): Single consonants

The main effects of Env and Place are significant. In general it takes longer for listeners to correctly identify a final consonant compared to an initial consonant (1492 vs. 1406 ms) (Env: (p = 0.045)). It also takes more time to recognize labials than coronals (1480 vs. 1419 ms) (Place: (p = 0.035)).[48] The factor of Pal is not significant, that is it takes about the same time to identify single plain and palatalized consonants.

There is a significant interaction of Place and Env (p = 0.035). A post-hoc test shows that it is due to the difference between the final labials and all other levels (initial labials (p = .009), initial coronals (p = .010) and final coronals (p = .026). It takes more than 150 ms longer to identify the stimuli with final labials than all other stimuli. Note that the RT for the final /p'/ is 256 ms longer than the value for the same consonant in the initial position (1521 vs. 1257 ms). In sum, the analysis of RT of single consonants shows sensitivity to Env and Place. This is mainly due to the longer time for the recognition of labials in codas, and particularly the final /p'/.

The factors that are significant for both identification and RT are Env and Place, as well as the interaction Place × Env.

6.1.4.2 Discussion

The results show that, with the exception of the final /p'/, the correct identification rate of single consonants is very high. It is almost 1.00 for all onset consonants. Recall that in this environment the gestures of the consonants are fully realized. Both CV and VC transitions, as

[48] This is due to the final /p/ and /p'/.

well as bursts, contribute to the accurate recovery of the gestures/features. As we expected, the identification rate in the coda is significantly lower (Env). Our predictions with regard to the different patterning of the palatalized labial are also confirmed (Env × Place). Recall that gestures in coda often undergo reduction. This is particularly true for the fronting and raising of Tongue Body during /p'/. The drop in its recognition is, however, more than would be expected based on the Tongue Body gesture reduction only. It seems that both the shift of the gesture to the onset of the consonant and a shorter burst typically found before vowels (no measurements were made) contribute to its poorer recognition. The lower identification of the final /t'/ (Env × Pal) is also expected based on phonotactic restrictions, yet lower identification does not seem to follow from its gestural magnitude (little articulatory reduction was found for some subjects). /t/ shows no sign of poorer performance in coda. Note that even though final consonants have transitions to the following vowel (CV), they do not have as much information about the consonant as the CV transitions of the onset consonant. The fact that the second vowel in the utterances was often more stressed than the first one could have also contributed to the better identification of the onset consonants.

The RT results also confirm our general expectations in terms of patterning of place and environment.

Let us now look at the confusion patterns of the final consonants (Tables 6.4-.17). As we would expect, the final palatalized /p'/ is largely confused with its plain counterpart, /p/ (32 tokens). In other cases it is perceived as palatalized /t'/ (20) and plain /t/ (10). When misidentified, the final palatalized /t'/ is perceived as its plain counterpart in almost all cases (15). Unlike these two consonants, the final plain /p/ is less often confused with the palatalized /p'/, but rather with the coronal /t/ (10). In other words, in the final coda environment palatalized consonants are more likely to be interpreted as their plain counterparts, while the reverse is not true.

To summarize, the results confirm our predictions about perception of single consonants in onset and coda positions.

6.1.5 Clusters: Onset (C_2 in VC_1C_2V)

6.1.5.1 Results

Table 6.12 shows the ANOVA results for the identification of onset consonants in clusters. Figure 6.3 presents the mean values for each of the onset consonants in various clusters.

Table 6.12 Correct identification: Onset consonants in clusters (statistics)

	F	p
Place	0.36	0.557
Pal	**18.66**	**0.000**
EnvPlace	0.64	0.433
EnvPal	**27**	**0.000**
Place × Pal	**6.62**	**0.019**
Place × EnvPlace	**10.14**	**0.004**
Place × EnvPal	**7.9**	**0.011**
Pal × EnvPal	**4.21**	**0.054**
Place × Pal × EnvPal	**36.79**	**0.000**

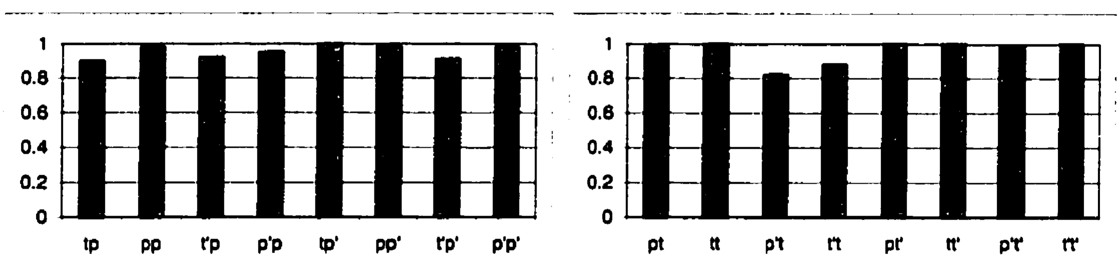


Figure 6.3 Correct identification: Onset consonants (C_2) in clusters

The main effects of EnvPal and Pal are significant. The effect of EnvPal, having a higher F value, shows that onset consonants are better identified when preceded by plain rather than palatalized consonants (.98 vs. .93) (p < .001) According to the effect of Pal, palatalized consonants are more often recognized than the plain ones (.98 vs. .93) (p < .001).

A number of factor interactions are significant. Of these the three-way interaction Place × Pal × EnvPal shows the strongest effect. It suggests that the perception of various onset consonants in clusters may be different depending on the secondary articulation of the preceding consonant. The post-hoc analysis shows that the major difference is between /t/ preceded by a palatalized consonant (C't: /p't/ and /t't/) and /t/ preceded by a plain segment (Ct) (.85 vs. .99). In fact the C'-t sequence is different from all other clusters (p = .000). Compare the rate of C't (.85) to all the other clusters (.95-.99).[49] Among the two C't clusters /p't/ has the lowest identification rate (.82). All other significant interactions seem to be due to this asymmetrical behaviour of /t/ after palatalized consonants.[50]

The ANOVA results for the RT of single consonants are given in Table 6.13. Figure 6.4 shows the mean RT values for each of the consonants in both environments.

Table 6.13 RT: Onset consonants in clusters (statistics)

	F	p
Place	**35.51**	**0.000**
Pal	**23.41**	**0.000**
EnvPlace	0.05	0.834
EnvPal	**21.96**	**0.000**
Place × EnvPlace	**7.08**	**0.015**
Place × EnvPal	**7.64**	**0.012**
Pal × EnvPal	4.21	0.054
Place × Pal × EnvPal	**5.85**	**0.026**

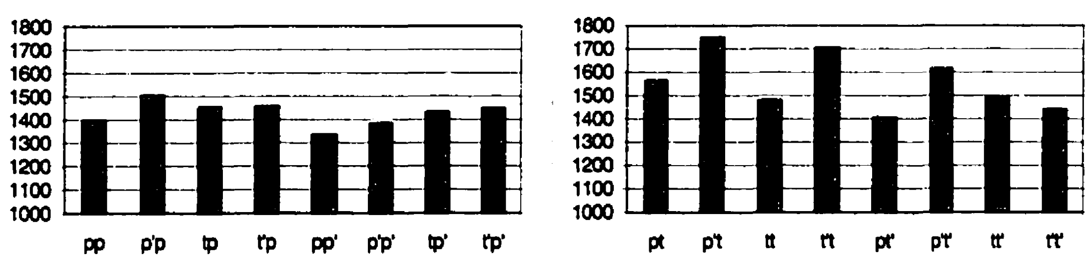


Figure 6.4 RT (ms): Onset consonants (C_2) in clusters

The results show the significance of EnvPal, Pal, and Place. EnvPlace is not significant. Identification of consonants preceded by palatalized segments is characterized by higher RT (1536 vs. 1442) (EnvPal: p < .001). It takes longer to identify plain than palatalized consonants

[49] Other significant differences are between various onset consonants (e.g. /t/ vs. /p/, /p'/ and /t'/ vs. /p'/ in certain environments), rather than between the same consonant in various environments.

[50] The interaction of Place and Pal (the Consonant factor) shows that the perception of /t'/ is significantly different from /p/ (p = .001) and /t/ (.000). Note that /t'/ is almost always correctly identified (1.00). There is also a difference between /t/ and /p'/ (p = .009). The perception of /p/ vs. /p'/ is not different. The identification of consonants stated in order of decrease is as follows: t' (1.00), p' (0.97), p (0.94), and t (0.92). The interaction of Place and Envpl is due to the difference between the combination labial-labial vs. coronal-labial (p = .039), where the identification is lower for the latter sequence (.97 vs. .93). The interaction of Place and EnvPal reveals that the sequence of a coronal

(1537 vs. 1441) (Pal: $p < .001$). It also takes more time to correctly perceive coronals than labials (1554 vs. 1424) (Place: $p < .001$).[51]

Some of the interactions are significant. The three-way interaction Place × Pal × EnvPal shows a significant difference between /t/ preceded by plain (Ct) and palatalized consonants (C't) (1520 vs. 1726). Note that the RT for /t/ is the highest when it is preceded by a hetero-organic palatalized consonant /p'/ (1747) and it is the lowest when it is preceded by the homorganic plain consonant /t/ (1477) (See Figure 6.4). Other interactions seem to be affected by the behaviour of /t/.[52]

The factors that are significant for both identification and RT factors are EnvPal and Pal, as well as the following interactions Place × EnvPlace, Place × EnvPal and Place × Pal × EnvPal.

6.1.5.2 Discussion

As we expected, the results show a fairly high identification rate of the onset consonants in clusters. This environment provides two crucial pieces of information about the quality of consonants: the spectral CV transition and burst. The first one allows for the recovery of the trajectories of primary and secondary articulators at the release of the consonant. The burst information is crucial to the palatalization and place distinctions.

While we did not expect a major influence of the preceding consonant, it turns out to be a factor. The rate of identification of C_2 is affected by C_1. This, however, is limited primarily to two particular clusters (/p't/ and /t't/) with a palatalized C_1 and plain /t/ as C_2. When misinterpreted in these clusters, /t/ is most often perceived as a palatalized /t'/ (14 out of 16 tokens). The values of Tongue Body of onset consonants in clusters were not examined in the

---

preceded by a palatalized consonant (C'-coronal) is less identifiable than when it is preceded by a plain consonant (C-coronal) (.92 vs. 1.00).

[51] The last two effect seem to be due mainly to the higher RT of /t/.

[52] The interaction Place × EnvPlace is due to the differences labial-labial vs. labial-coronal and coronal-coronal ($p = .001$), as well as coronal-labial vs. labial-coronal ($p = .000$) and coronal-coronal ($p = .028$). ordered labial-labial (1402), coronal-labial (1446), coronal-coronal (1527), labial-coronal (1581). The interaction Place × EnvPal involves a major difference between the sequence C'-coronal and all other clusters (C-labial C'-labial, and C'-coronal) ($p = .000$). C'-coronal has the highest RT (1625), while the other sequences are in the range (1401-1483).

articulatory study. It is reasonable to expect that the Tongue Body value of /t/ is higher after a palatalized gesture and thus easily confused with /t'/. This may be of less importance for clusters /p'p/ and /t'p/ due to the longer duration of /p/. Also note that in both of these clusters C_1 is rarely released (the effect of place of articulation order in /p't/ and homorganicity in /t't/). Note also that neither /p't/, nor /t't/ are possible within words in Russian, and thus listeners may be less familiar with these sequences.

The overall higher identification and lower RT for /t'/ compared to /t/ is consistent with my previous findings with final /t'/ and /t/ (see Chapter 4).

6.1.6 Clusters: Coda (C_1 in VC_1C_2V)

6.1.6.1 Results

Table 6.14 shows the ANOVA results for the identification of coda consonants in clusters. Figure 6.5 presents the mean values for each of the consonants in various clusters.

Table 6.14 Correct identification: Coda consonants in clusters (statistics)

	F	p
Place	3.5	0.077
Pal	1.04	0.321
EnvPlace	1.16	0.294
EnvPal	**15.51**	**0.000**
Place x Pal	**100.31**	**0.000**
Place x EnvPlace	**16.47**	**0.000**
Pal x EnvPlace	**21.29**	**0.000**
Pal x EnvPal	**205.34**	**0.000**
EnvPlace x EnvPal	**7.66**	**0.012**
Place x Pal x EnvPlace x EnvPal	**36.66**	**0.000**

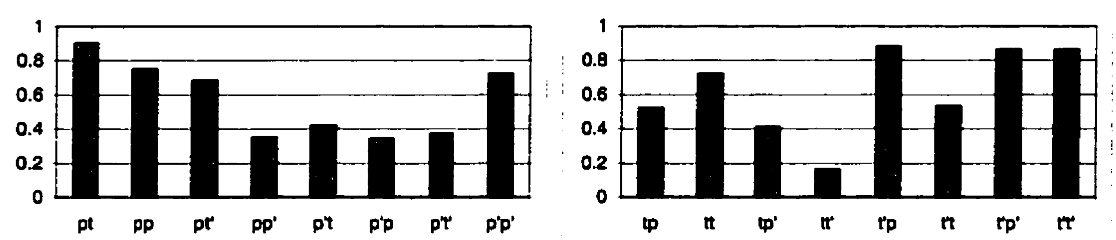


Figure 6.5 Correct identification: Coda consonants (C_1) in clusters

Of all the main effects only EnvPal is significant, while EnvPlace, Pal, and Place are not significant. Consonants are better identified before plain rather than palatalized consonants (.63 vs. .55) (EnvPal: ($p < .001$)).

There are a number of factor interactions. Among these the interaction Pal-EnvPal and Place × Pal are of particular magnitude (based on F values). The interaction Pal × EnvPal indicates the relation between the secondary articulations of C_1 and C_2. The analysis shows that perception of sequences where two consonants differ with respect to palatalization (CC' and C'C) is significantly different from all other combinations ($p = .000$). Their correct identification is the lowest (0.4 and 0.54), while the consonants in clusters homogeneous with respect to palatalization show the highest identification (0.72 and 0.90). The Place × Pal interaction (or the Consonant factor) indicates that both /t'/ and /p/ are significantly different from all other consonants (all differences $p = .000$, except for /t'/ vs. /p/ ($p = .036$)). Both /t'/ and /p/ are the consonants most reliably identified (0.78 and 0.67), while /p'/ and /t/ are most often misidentified (0.46 and 0.45). The interaction Pal × EnvPlace reveals that palatalized consonants are more easily identified than plain segments before labials (C'-labial vs. C-labial ($p = .001$). It also shows that palatalized segments are better perceived before labials than before coronals (C'-labial vs. C'-coronal ($p = .006$). The interaction EnvPlace × EnvPal shows that the environment before /t'/ is different from all other contexts (/p/ ($p = .000$), (/p'/ ($p = .032$), (/t/ ($p = .000$)). The identification in this environment is the lowest (.52) compared to the other contexts (.59-.64). The four-way interaction (Place × Pal × EnvPlace × EnvPal) is also significant. Thus both the quality of C_1 and C_2 contribute to the identification rate of a particular coda consonant.[53] In sum, there are a number of significant interactions due to asymmetries in perception of the four consonants in various environments.

[53] There are few interesting significant differences grouped by the coda consonant: plain labials (pp' vs. pt ($p = .000$), pt' ($p = .038$), and pp (.007)), palatalized labials (p'p' vs. p'p ($p = .012$), p't' ($p = .026$), p'p vs. p't' ($p = .000$), plain coronals (tp vs. tt' ($p = .018$) and tt vs. tt' ($p = .000$)), and palatalized coronals (t'p vs. t'p' ($p = .000$), t't ($p = .023$), t't' ($p = .000$), t't vs. t'p ($p = .023$), t'p' ($p = .038$), t't' ($p = .038$), t'p' vs. t't' ($p = .000$)).

The ANOVA results for the RT of single consonants are given in Table 6.15. Figure 6.6 shows the mean RT values for each of the consonants in both environments.

Table 6.15 RT: Coda consonants in clusters (statistics)

	F	p
Place	2.30	0.227
Pal	.217	0.673
EnvPlace	1.13	0.366
EnvPal	2.50	0.212
Pal × EnvPlace	**10.08**	**0.050**
Pal × EnvPal	**122.92**	**0.002**

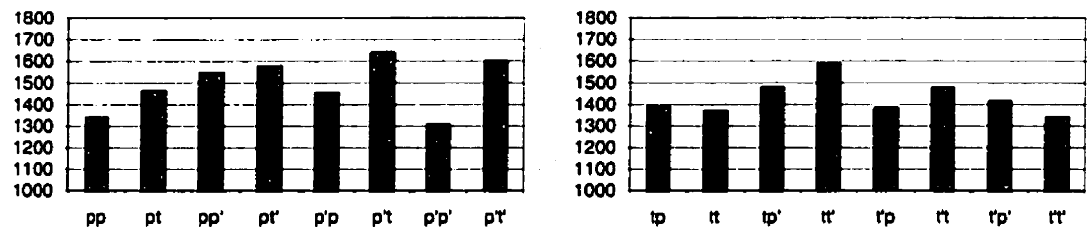


Figure 6.6 RT (ms): Coda consonants (C_1) in clusters

None of the main effects (EnvPal, EnvPlace, Pal, and Place) are significant. There is, however, a significant interaction between Pal × EnvPal and Pal × EnvPlace. The effect of Pal × EnvPal, as in the case with correct identification, is highly significant. The analysis demonstrates that plain-plain clusters (CC) exhibit significantly lower RTs than all other combinations (CC' (p = .003, C'C (p = .006), C'C' (p = .017). RT for the plain-palatalized sequence (CC') is significantly higher for the palatalized-palatalized cluster (C'C' (p = .034). The following are all cluster types ordered in terms of their RT values (from lowest to highest): CC (1390 ms), C'C' (1414 ms), C'C (1486 ms), and CC' (1547 ms). The interaction Pal × EnvPlace shows that it takes longer to recognize a palatalized consonant before a coronal (C'-coronal) than before a labial (C'-labial) (1512 vs. 1388 ms) (p = .041). Plain consonants do not show the effect of the following place of articulation.

The interactions that are significant for both identification and RT are Pal × EnvPal and Pal × EnvPlace.

6.1.6.2 Discussion

The results confirm a number of our predictions. The overall identification in preconsonantal coda position is relatively low. This environment has less information about the primary and, particularly, the secondary gesture/feature of C_1 (due to TB reduction in coda and influence of C_2 due to overlap). We may not exclude the possibility that the lower degree of stress on V_1 may have also contributed to the poorer identification of C_1.

As we expected, the identification of palatalized consonants is lower before plain consonants (Pal × EnvPal), and particularly before plain homorganic consonants (Pal × EnvPal × EnvPlace). In the latter position there is a tendency for greater reduction of TB-palatal and less frequent release of C_1 that provides the valuable acoustic information about the contrast. The rate of identification of /t'/ before plain consonants, however, is still fairly high (.53 in _t and .88 in _p). It is also surprising that /t/ before /p/ is rather often perceived as /t'/ (29 tokens). It is possible that this is due to the influence of stem-internal phonotactics: the [d'b] sequence ([t'p] is absent) is more common within stems, while [tp] and [db] occur across prefix-stem boundaries (see section 2.2.2.2.1.3).

As was predicted, we find that plain consonants are more likely to be confused before palatalized consonants, and particularly, if the latter are of the same place of articulation. Thus, the degree of confusion is the greatest for /tt'/, followed by /pp'/ and then /tp'/. This correlates with our articulatory findings (infrequent release of C_1; TB values of C_1 are greatly influenced by C_2; see section 5.5.2.4) and the assimilation patterns found in Russian and its dialects (Avanesov 1972, Avanesov & Orlova 1965; see section 2.2.2.2.2). Overall, the identification of coda consonants in sequences that disagree with respect to secondary articulation are more problematic. The RT results provide additional evidence for this.

The expected differences with respect to the place of articulation of the palatalized consonants are well supported by our results. The palatalized labial /p'/ shows a high error rate in the preconsonantal environment. It is often confused with its plain counterpart /p/ (49 cases in _t, 42 in _p, and 41 in _t'). This is due to both its articulatory (lower TB, more overlap with the following coronal) and acoustic (less salient burst if it is released) properties.

Note that there is also a tendency towards perceptual place assimilation: the place of C_1 is sometimes confused with the place of C_2. Thus, /t/ is confused with /p/ in the environments _p (15) and _p' (16) and /p/ is misperceived as /t/ in the contexts _t (9), _t' (5). As in the case of palatalization, this kind of perceptual confusion is likely to be a motivating factor behind place assimilation. Notice also that the coronal /t/ has a greater tendency to assimilation than /p/.

In sum, the results meet our expectations about the perception of coda consonants and the interactions of environment, place, and palatalization.

6.1.7 Single vs. Clusters

In this section I compare the similarity of identification patterns and RT between single consonants and consonants in clusters combined in terms of syllable position. Thus, there are two groups: onset consonants and coda consonants.

6.1.7.1 Results

The ANOVA[54] results for identification showed that there was no significant difference between single and cluster consonants in the onset position (the Factor of Type) [F = 3.049; p = .106]. Neither the main effects of Place [F = 2.59; p = 134] and Pal [F = 3.251; p = .097], nor any interactions were significant. The same results were found for the coda consonants: no difference in terms of Type [F = 1.317; p = .273], Place [F = .046; p = .834], and Pal [F = 1.119; p = .311]. No significant interactions were found.

The ANOVA results for RT also suggest the similarity between the two types of consonants. None of the factors was significant either for the onset (Type [F = 1.802; p = .204], Place [F = 2.543; p = .137], Pal [F = .923; p = .356]), or for the coda consonants (Type [F = .299; p = .594], Place [F = 2.922; p = .113], Pal [F = .178; p = .681]).

6.1.7.2 Discussion

These results show that in general perception of the consonants in question is more dependent on their major environments: syllable onset (or prevocalic) and syllable coda (or post-vocalic) than on whether the consonants are single or in clusters. This is particularly true for the onset consonants: the distribution patterns of palatalization are rarely dependent on whether the consonants we are dealing are single consonants or consonants in clusters (see section 2.1.3.2.3). In general the preconsonantal coda consonants do show lower identification rates than the final coda consonants. But, apparently, since the factors show effects in the same direction, there are no statistical differences between the two groups.

6.1.8 General discussion

In this section I review the results by looking first at particular consonants and then at environment in general.

6.1.8.1 Consonants

Which consonants are the most likely and which are the least likely to be confused in general and with respect to their secondary articulation?

Figure 6.7 shows the overall identification rate for the four stops (averaged by environment). An important difference between them is that plain labial /p/ and palatalized coronal /t'/ are more often correctly recognized than their counterparts in terms of palatalization, /p'/ and /t/.

[54] Total mean values for all utterances averaged across listeners were evaluated in an ANOVA with three factors: Type (single or cluster), Place (labial vs. coronal), and Pal (plain vs. palatalized), as described above.

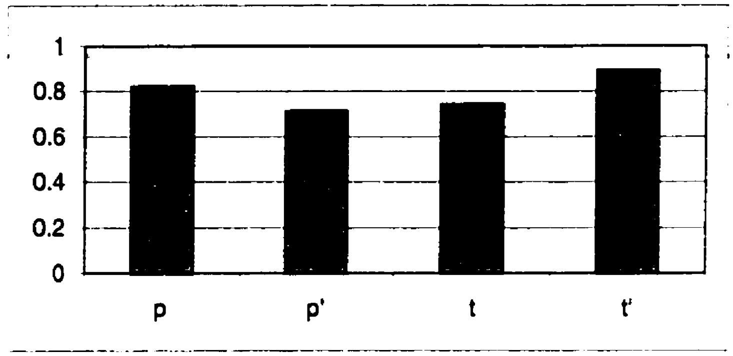
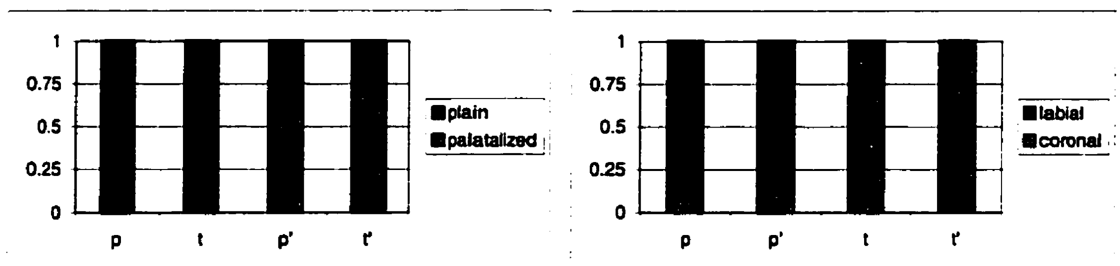


Figure 6.7 Mean correct consonant identification

If misidentified, plain and palatalized consonants of different places of articulation are most likely to be confused with each other than with consonants of the other place (see Tables 6.4-6.7). There are, however, asymmetries in the confusability patterns. /p'/ is perceived as /p/ 19% of the time (overall), while /p/ is identified as /p'/ only in 10% of the cases. The difference is greater for coronals: /t/ is labeled as /t'/ in 16% of cases, while the reverse is true only 5% of the time.

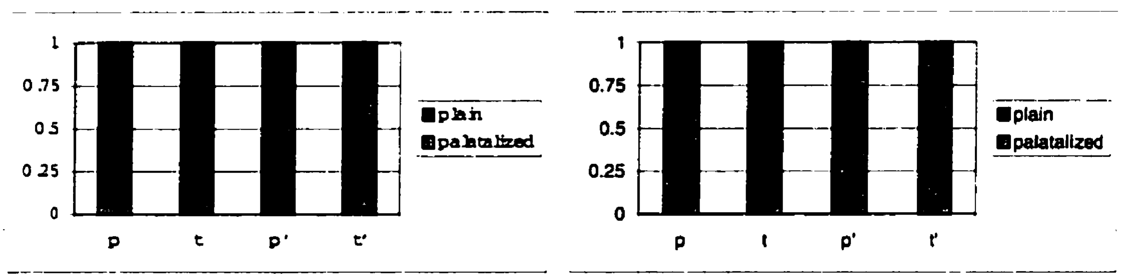
Looking at identification and confusability in terms of palatalization and place (see the right columns in Tables 6.4-6.7) allows us to determine to what extent each of the consonants is perceived as plain or palatalized and as labial or coronal. Figure 6.8 shows that /p/ is predominantly identified as plain (.87) and labial (.93), while /t'/ is overwhelmingly perceived as palatalized (.95) and coronal (.92). The responses for /p'/ and /t/ are less consistent than for /p/ and /t'/ both in terms of palatalization and place, but identification is still well above the chance level (.50). The consonants tend to show less variation in place than in palatalization.



a. b.

Figure 6.8 Mean perception of the consonants in terms of palatalization (a) and place (b)

These results are averaged for environment. As we know, the perception of consonants is sensitive to syllable position and the quality of neighbouring segment. Figure 6.9 illustrates the perception of the consonants in term of palatalization in two of the 'worst' environments, before plain (a) and palatalized (b) homorganic consonants.



a. b.

Figure 6.9 Perception of the consonants in terms of palatalization before plain (a) and palatalized (b) homorganic consonants

We can see that, compared to Figure 6.9a, the palatalized consonants before a plain homorganic segment, and especially /p'/ (at the chance level), are more often perceived as plain. The difference in the perception of plain consonants before homorganic palatalized segments is particularly substantial (Figure 6.9b). Note that /t/ is perceived to have more palatalized quality in this context than /p/. It appears that the tendency towards perceptual 'assimilation' is stronger before palatalized consonants than the plain ones. The degree of 'assimilation' of C_1 to the place of C_2 is less common in our results.

6.1.8.2 Environments

Let us now turn to the environments. Based on our results, what are the 'best' and the 'worst' environments for correct identification of stops and for perception of the plain-palatalized contrast? Figure 6.10 presents the overall identification rate by environment (averaged by consonants), ordered from 'best' to 'worst'.

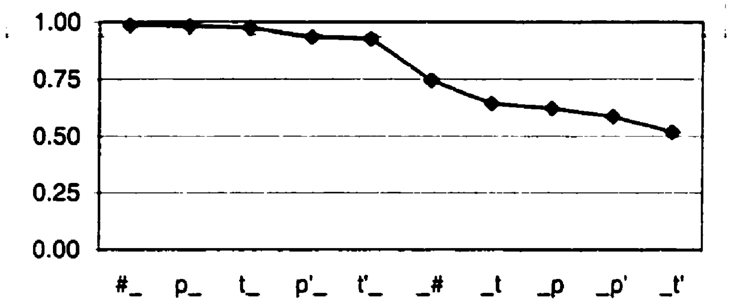


Figure 6.10 Overall identification of the consonants in all environments

We can see that identification is most successful in the onset position, especially word-initial or after plain consonants (close to 1.00). It is somewhat lower after palatalized segments. The recognition of coda consonants is much lower, however, it varies from context to context. The word-final environment shows a moderate drop in identification, followed by the contexts before plain and then palatalized consonants (almost .50). Note that Figure 6.10 presents the general correct identification. Some of the errors made are in terms of secondary articulation, while others are in terms of place of articulation. For our purposes the secondary articulation errors are of particular importance. It would be useful for our analysis to differentiate these errors from the place confusions.

Figure 6.11 shows the correct identification split by correct recognition of the plain-palatalized and labial-coronal contrasts.

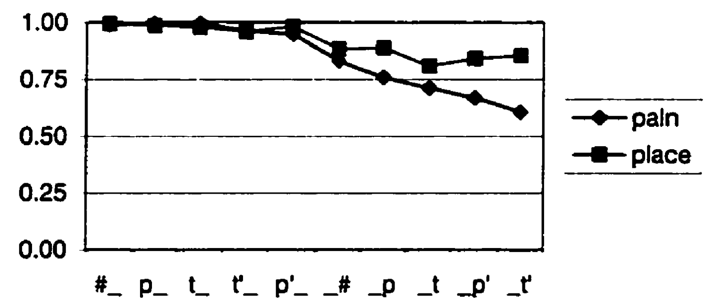


Figure 6.11 Identification of consonants
in terms of palatalization (paln) and place in all environments

We can see that both palatalization and place errors occur predominantly in the coda position, however, palatalization is obviously more affected than place. Place identification does not seem to show a clear dependence on whether the consonant is single or is followed by another segment and on what the quality the following consonant is. On the contrary, the recognition of palatalization is lower for segments in clusters than for single consonants. On the other hand, it is higher when the coda consonant is followed by a palatalized segment, and it is lower before coronal (both plain and palatalized) consonants than before labials.

6.1.8.3 Consonants and environments

Now we will look at the difference between identification of consonants in terms of palatalization in particular environments. Figure 6.12 presents the overall results for plain (a) and palatalized (b) consonants.

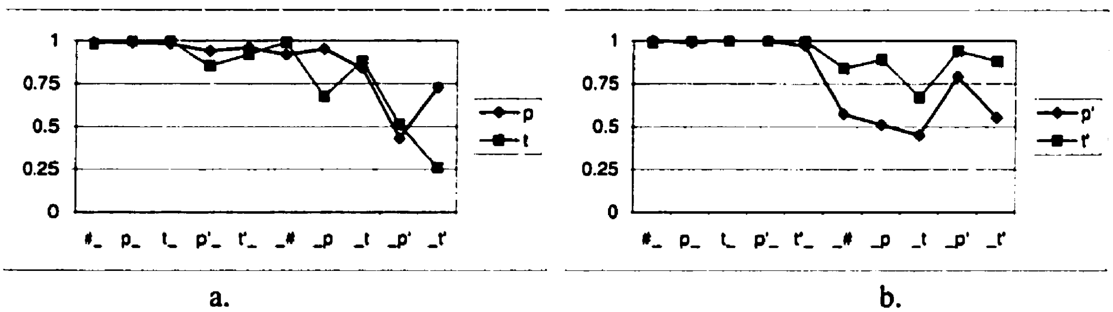


Figure 6.12 Identification of consonants in terms of palatalization in all environments: plain (a) and palatalized (b) stops

There is a general similarity between both plain and palatalized consonants, and there are also some differences. For instance, the identification of /t/ is very similar to /p/ with the exception of two contexts, before /p/ and before /t'/. While the plain labial (Figure 6.12a) shows a high rate of identification in these environments, the plain coronal is more often confused with its palatalized counterpart. Thus the factor of hetero-/homorganicity affects the two consonants in different ways.

/p'/ and /t'/ show similar patterns but differ in the rate of identification in coda (Figure 6.12b). While the homorganicity of the following plain consonant is important for /t'/, it is not for /p'/. At the same time, /p'/ is sensitive to whether the following palatalized consonant is homorganic or not.

The perception of palatalized consonants seems to be shaped by syllabic position to a greater extent than by specific environments. Plain consonants show less difference between the general onset and coda environments, rather showing more sensitivity to specific contexts.

The two most salient consonants, /p/ and /t'/, seem to define two extremes: /p/ is a prototypical plain and labial consonant, while /t'/ is the most characteristic palatalized and coronal consonant. The other two consonants are in between.

In terms of the direction of influence, we find that it is a following consonant that influences the perception of a preceding consonant rather than vice versa. The only example of progressive influence was the effect of a palatalized consonant on the following /t/. This influence, however, is less strong than the effects of the following palatalized segments.

6.1.8.4 Conclusion

The results of perception of stops in a variety of environments by Russian listeners allowed us to determine that perception of the same consonant is very different depending on the environments and the consonants themselves. The relation between the asymmetries found in identification and RT to the articulatory factors will be discussed in Chapter 7.

In the next section I investigate the perception of the same stimuli in noise. This will allow us to determine the robustness of certain characteristics of the stops and compare their recognition with the 'ideal' situation presented in this section.

6.2 Identification task (noise): Russian listeners

6.2.0 Introduction

It is reasonable to expect that the overall rate of identification of a sound will be lower in noise than under ideal conditions discussed in section 6.1. Some of the important acoustic information necessary for recovering gestures/features is masked by the noise. This is particularly true for stop bursts that are characterized by certain noise frequencies. The vocalic transitions are known to be more noise-resistant than burst (Liberman 1996). Given this, confusion is likely to be made between gestures whose acoustic patterns (for instance F2 values) are closer together. Thus, /t'/ and /p'/ would be harder to distinguish in noise due to masking of their bursts and very similar CV transitions (due to the higher TB gesture). /p'/ and /t/ are also more likely to be confused than /t'/ and /p/, whose CV and VC transitions are very far apart (due to the difference in TB position and the primary gestures, Lips and Tongue Tip). More importantly, we can expect that some of the effects found in the previous sections will show up more robustly under adverse condition.

6.2.1 Experimental setup, materials and procedure

The same twenty native speakers of Russian participated in this experiment. The general setup and procedure were the same as in the experiment with clean stimuli (section 6.1.1). A new set of acoustic stimuli was also prepared: noise was added to the tokens from the previous experiment (nonsense phrases of the type *ta*C_1*#* C_2*apy)* by using SoundEdit (option 'Add Noise'). The onset and offset of noise were tapered to avoid an audible click. Each token had a duration of about 600 ms. The level of noise, determined in a pilot study with three subjects, was considered to be adequate to induce a large number of errors, while maintaining recognition above the chance level. The same stimuli were used for onset and coda consonants in clusters.

6.2.3 Analysis

The analysis was the same as in the previous experiment (see section 6.1.2). Due to the high error rate in the noisy data, the tokens were not examined for response time (RT). Recall that within each group I ran separate repeated-measures ANOVAs for three sub-groups: single consonants (onset and coda together), onset consonants in clusters, and coda consonants in clusters. This division of the data allowed me to examine the full interaction of factors. The within-subject factors used in the ANOVAs for single consonants were Place (2 levels: labial vs. coronal), Palatalization (*Pal,* 2 levels: plain vs. palatalized), and Environment (*Env*, 2 levels: onset vs. coda). The factors used for consonants in clusters (onset and coda) were Place (2 levels: labial vs. coronal), Palatalization (*Pal*, 2 levels: plain vs. palatalized), Environment Place (*EnvPlace* (2 levels: labial vs. coronal), and Environment Palatalization (*EnvPal*, 2 levels: plain vs. palatalized). In both analyses the dependent variable corresponded to 20 repeated measurements of identification per subject. I used a Tukey HSD post-hoc test to compare means in significant interactions. In order to compare the results for single consonants and consonants in clusters, the total mean values for each utterance were analyzed in ANOVAs for onset and coda environments. The factors were Type (2 levels: single vs. cluster), Place (2 levels: labial vs. coronal), and Pal (2 levels: plain vs. palatalized). The same analysis was performed for both correct identification and RT.

6.2.4 Overall results

In this section I give a general overview of the results of the perception study under conditions of noise. Means and standard deviations for correct identifications of the four stops, based on the average for all subjects, are presented in Tables 6.16-6.17. The stimuli are divided into groups by syllable position. Identification matrices for /p/, /p'/, /t/, and /t'/ are shown in Tables 6.18-6.21.

Table 6.16 Correct identification: Single consonants (means and standard deviations)

initial	*Mean*	*SD*	*final*	*Mean*	*SD*
ta papy	0.73	0.20	tap apy	0.75	0.21
ta p'apy	0.25	0.24	tap' apy	0.23	0.22
ta tapy	0.29	0.19	tat apy	0.14	0.13
ta t'apy	0.77	0.29	tat' apy	0.12	0.19
	0.51	*0.23*		*0.31*	*0.19*

Table 6.17 Correct identification: Consonants in clusters (means and standard deviations)

onset	*Mean*	*SD*	*coda*	*Mean*	*SD*
tat papy	0.52	0.29	tap tapy	0.55	0.32
tap papy	0.64	0.26	tap papy	0.52	0.20
tat' papy	0.38	0.21	tap t'apy	0.53	0.27
tap' papy	0.52	0.28	tap p'apy	0.48	0.25
tat p'apy	0.12	0.14	tap' tapy	0.23	0.21
tap p'apy	0.18	0.17	tap' papy	0.14	0.16
tat' p'apy	0.16	0.17	tap' t'apy	0.21	0.24
tap' p'apy	0.13	0.16	tap' p'apy	0.19	0.18
tap tapy	0.32	0.24	tat papy	0.21	0.20
tat tapy	0.28	0.34	tat tapy	0.11	0.14
tap' tapy	0.31	0.24	tat p'apy	0.13	0.15
tat' tapy	0.1	0.15	tat t'apy	0.18	0.17
tap t'apy	0.76	0.23	tat' papy	0.19	0.22
tat t'apy	0.71	0.30	tat' tapy	0.25	0.25
tap' t'apy	0.66	0.23	tat' p'apy	0.27	0.27
tat' t'apy	0.74	0.21	tat' t'apy	0.18	0.18
	0.41	*0.23*		*0.27*	*0.21*

The total average identification rate was .35 (.22) (compare to .79 (.15) with clean stimuli). The rate for all single consonants was 0.41 (0.21) and for consonants in clusters was .35 (.22). Within the single consonants a rate of .51 (.23) was found for the onset (initial) and a rate of .31 (.19) for the coda (final). In clusters the onset consonants had a rate of .41 (.23) and the coda segments a rate of .27 (.21). Collapsing all the groups into onset and coda classes, the total rate for the onset is .43 (.23), and the total rate for the coda is .28 (.21).

Table 6.18 Identification matrix: /p/ (tokens)

		Consonant				Place		Palatalization	
		p	p'	t	t'	**lab**	cor	**npal**	pal
single	a._a	**74**	3	18	4	**77**	22	**92**	7
	a_.a	**74**	16	5	4	**90**	9	**79**	20
CC-onset	t_a	**50**	10	22	16	**60**	38	**72**	26
	p_a	**64**	6	11	19	**70**	30	**75**	25
	t'_a	**39**	9	29	21	48	**50**	**68**	30
	p'_a	**55**	6	27	11	**61**	38	**82**	17
CC-coda	a_t	**56**	19	17	8	**75**	25	**73**	27
	a_p	**51**	28	15	5	**79**	20	**66**	33
	a_t'	**53**	19	14	13	**72**	27	**67**	32
	a_p'	**50**	17	21	12	**67**	33	**71**	29
	total	***566***	*133*	*179*	*113*	***699***	*292*	***745***	*246*
	ratio	***0.57***	*0.13*	*0.18*	*0.11*	***0.71***	*0.29*	***0.75***	*0.25*

Table 6.19 Identification matrix: /p'/ (tokens)

		Consonant				Place		Palatalization	
		p	p'	t	t'	lab	cor	npal	pal
single	a._a	8	26	7	**59**	34	**66**	15	**85**
	a_.a	**49**	23	9	19	**72**	28	**58**	42
CC-onset	t'_a	6	15	7	**72**	21	**79**	13	**87**
	p'_a	9	10	28	**53**	19	**81**	37	**63**
	t_a	8	11	13	**68**	19	**81**	21	**79**
	p_a	14	18	11	**56**	32	**67**	25	**74**
CC-coda	a_t	**52**	23	9	16	**75**	25	**61**	39
	a_p	**50**	14	21	14	**64**	35	**71**	28
	a_t'	**58**	21	10	9	**79**	19	**68**	30
	a_p'	**54**	18	12	16	**72**	28	**66**	34
	total	*308*	*179*	*127*	***382***	*487*	***509***	*435*	***561***
	ratio	*0.31*	*0.18*	*0.13*	***0.38***	*0.49*	***0.51***	*0.44*	***0.56***

Table 6.20 Identification matrix: /t/ (tokens)

		Consonant				Place		Palatalization	
		p	p'	t	t'	lab	cor	npal	pal
single	a._a	**42**	10	28	19	**52**	47	**70**	29
	a_.a	**57**	13	14	16	**70**	30	**71**	29
CC-onset	p_a	**56**	1	32	11	**57**	43	**88**	12
	t_a	**35**	12	26	26	47	**52**	**61**	38
	p'_a	17	16	31	**36**	33	**67**	48	**52**
	t'_a	28	12	9	**51**	40	**60**	37	**63**
CC-coda	a_p	**37**	20	21	22	**57**	43	**58**	42
	a_t	**59**	17	11	13	**76**	24	**70**	30
	a_p'	**42**	22	13	23	**64**	36	**55**	45
	a_t'	**41**	18	18	23	**59**	41	**59**	41
	total	***414***	*141*	*203*	*240*	***555***	*443*	***617***	*381*
	ratio	***0.41***	*0.14*	*0.20*	*0.24*	***0.56***	*0.44*	***0.62***	*0.38*

Table 6.21 Identification matrix: /t'/ (tokens)

		Consonant				Place		Palatalization	
		p	p'	t	t'	lab	cor	npal	pal
single	a._a	4	10	8	**78**	14	**86**	12	**88**
	a_.a	**59**	12	17	12	**71**	29	**76**	24
CC-onset	p'_a	12	9	13	**65**	21	**78**	25	**74**
	t'_a	3	15	10	**72**	18	**82**	13	**87**
	p_a	6	8	11	**75**	14	**86**	17	**83**
	t_a	4	15	10	**71**	19	**81**	14	**86**
CC-coda	a_p	**42**	18	21	19	**60**	40	**63**	37
	a_t	**41**	17	16	25	**58**	41	**57**	42
	a_p'	**41**	17	14	27	**58**	41	**55**	44
	a_t'	**61**	12	9	18	**73**	27	**70**	30
	total	*273*	*133*	*129*	***462***	*406*	***591***	*402*	***595***
	ratio	*0.27*	*0.13*	*0.13*	***0.46***	*0.41*	***0.59***	*0.40*	***0.60***

The overall RT rate in noise is 1680 (275) (compared to 1464 (237) ms with clean stimuli). It takes about 1660 (272) ms to correctly identify an onset consonant and 1701 (278) ms to recognize a coda segment. Since the identification rate in noise was relatively low, I did not perform a statistical analysis of the RT results.

6.2.5 Single consonants

6.2.6.1 Results

Table 6.22 shows the ANOVA results for the identification of onset consonants in clusters. Figure 6.13 presents the mean values for each of the onset consonants in various clusters.

Table 6.22 Correct identification: Single consonants (statistics)

	F	p
Place	**25.71**	**0.000**
Pal	**19.77**	**0.000**
Env	**45.06**	**0.000**
Place × Pal	**97.2**	**0.000**
Place × Env	**35.12**	**0.000**
Pal × Env	**15.06**	**0.001**
Place × Pal × Env	**12.62**	**0.002**

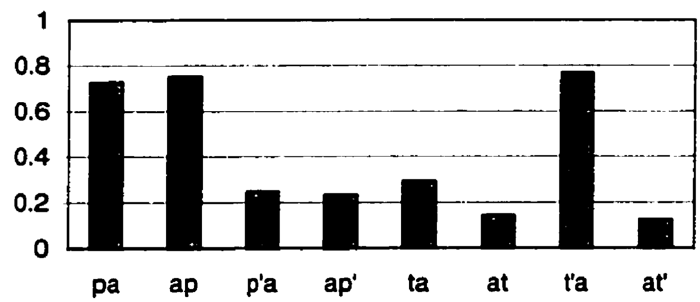


Figure 6.13 Correct identification: Single consonants

All of the factors, Place, Pal, and Env, are significant. As expected, Env is the most important factor: it has the highest F value of all the factors and interactions. Correct identification of consonants in coda was significantly lower than in onset (.31 vs. .51) (Env: [$p < .001$]). Note that the former is only slightly above the chance level (.25). Palatalized consonants were less correctly identified than their plain counterparts (.34 vs. .48) (Pal: [$p < .001$]). The third factor was Place: coronals were less reliably identified than labials (.33 vs. .49) (Place: [$p < .001$]).

All the interactions of the factors are significant. The interaction Place × Pal (the largest among all factors and interactions) shows that both /p/ and /t/ are different from each other and all other consonants (/p/ vs. all .000; /t'/ vs. all, /p'/ .005, /t/ .002) ($p = .000$). These two, and particularly /p/, are more perceptually robust in noise. The recognition of /p'/ and /t/ is below chance level.

The interaction Place × Env indicates that perception of single consonants is sensitive to the place of articulation in a given environment. A Tukey HSD post-hoc test determined that only final coronals are different from all other levels of factors ($p = .000$). Their perception is much worse (.13) compared to the identification of initial coronals (.53), initial and final labials (both .49).

According to the analysis of the Pal × Env interaction, the identification rate of the final palatalized consonants (.18) is significantly lower than the rate of initial palatalized consonants (.51) as well as initial and final plain consonants (.51 and .45) (all $p = .000$). No other differences are found.

The three-way interaction (Place × Pal × Env) is also significant. The post-hoc test determined that within consonants only initial and final variants of /t'/ are significantly different from each other ($p = .000$). Other differences involve /p/: the final /p/ is different from all final consonants ($p = .000$); the initial /p/ differs from /p'/ and /t/ ($p = .000$).

In sum, all the factors and all interactions of these factors are significant. This is mostly due to better recognition of /p/ and /t'/ and their different behaviour in the two syllabic positions.

6.2.6.2 Discussion

Recall that the study of clean stimuli showed that all consonants in onset position exhibit a very high identification rate. Perception of the same consonants in coda was lower, particularly for /p'/.

The results show that noise substantially affects the perception of stops. However, not all of the stops are affected to the same degree. As in clean conditions, we find asymmetric relations between place and palatalization. Plain labials are identified better than their palatalized counterparts under both conditions. The opposite is true of coronals: the palatalized /t'/ is significantly better identified plain /t/ (in onset), also under both conditions. As expected, these differences, quite subtle in the clean stimuli, become very substantial in noise.

The effect of environment is of particular interest to us. While two out of four consonants are reliably identified in the onset position (/p/ and /t'/), only one is recognized in the coda position (/p/). All other segments are recognized at or below the chance level. Recall that this effect was also present in the clean stimuli, but it affected mostly /p'/ and, to a lesser extent, /t/ and /t'/.

Looking at the confusion patterns of the onset consonants (Tables 6.18-.6.21) we can conclude that perception is very different from consonant to consonant in some respects, and similar in others. When misidentified, the initial /p/ is more often confused in terms of place (perceived as /t/ in 18 cases), while the final /p/ is more often misrepresented in terms of palatalization (perceived as /p'/ in 16 cases). Most of the time the initial /p'/ is perceived as a palatalized coronal /t'/ (59), while the same segment in the final position is identified as plain labial /p/. Correct identification of the segment (as /p'/) is the second choice in both of these environments (26 and 23). Both initial and final /t/ are most likely to be confused with plain labial /p/ (42 and 57). When presented with a /t'/ in an onset, listeners correctly identify it in the majority of the cases (78; /p'/ is the second choice: 10). The same consonant in the coda is perceived mostly as /p/ (59 cases) (with /t/ being the second choice: 17).

While both plain and palatalized consonants are about equally preferred in an onset, it is overwhelmingly the plain consonant that is chosen in the coda. In terms of place of articulation, a consonant in the coda is most likely to be perceived as a labial. Plain consonants in noise are much less often confused with the palatalized segments, while the palatalized consonants are predominantly identified as plain in the coda.

To summarize, the results confirm our prediction that the effects found in the study in the clean stimuli will be more robust under conditions of noise. This is particularly true for the effect of environment: the final position is much 'worse' for the palatalized segments. The asymmetry between /t'/ and /p'/ is also confirmed. The latter is very poorly identified, being confused either with /t'/ in the onset or with /p/ in the coda.

It appears that the listeners are able to reliably distinguish four onset contrasts and three coda contrasts under ideal conditions. When noise is added, their performance is much worse: they perceive only two contrasts in the onset and one in the coda.

6.2.7 Clusters: Onset (C_2 in VC_1C_2V)

6.2.7.1 Results

Table 6.23 shows the ANOVA results for the identification of onset consonants in clusters. Figure 6.14 presents the mean values for each of the onset consonants in various clusters.

Table 6.23 Correct identification: Onset consonants in clusters (statistics)

	F	p
Place	**16.91**	**0.001**
Pal	2.51	0.130
EnvPlace	**11.13**	**0.003**
EnvPal	**20.89**	**0.000**
Place x Pal	**72.57**	**0.000**
Pal x EnvPlace	**11.32**	**0.003**
Pal x EnvPal	**6.31**	**0.021**
Pal x EnvPlace x EnvPal	**10.41**	**0.004**

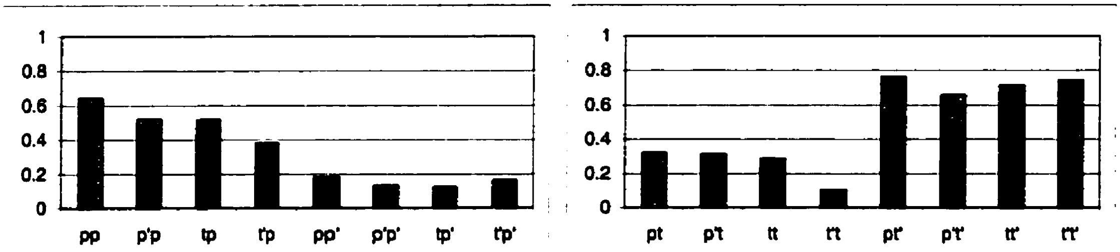


Figure 6.14 Correct identification: Onset consonants (C_2) in clusters

The main effects of EnvPal, Place, and EnvPlace are significant. The effect of EnvPal shows that onset consonants are better identified when preceded by plain rather than palatalized consonants (.44 vs. .37) ($p < .001$). In this environment coronals are more often recognized than labials (.49 vs. .33) (Place: $p < .01$). The environment after a labial is better for phoneme identification than the context after a coronal (.44 vs. .38) (EnvPlace: $p < .01$).

A number of factor interactions are significant. The interaction Place × Pal indicates the difference between the consonants. The post-hoc analysis shows that the effect is due to two consonants: /p/ and /t'/. They are different from each other ($p = .038$) and the other consonants (/p/ vs. /p'/ ($p = .000$), /t/ ($p = .006$); /t'/ vs. /p'/ ($p = .000$), /t/ ($p = .000$)). These two consonants are much better recognized (.51 and .72) than /p'/ and /t/ (.15 and .25).

The interaction Pal × EnvPlace is due to the difference between the sequence coronal-C and all other sequences (labial-C ($p = .001$), labial-C' ($p = .003$), coronal-C' ($p = .003$). No other differences are significant. The identification of a plain consonant is worse when it is preceded by a coronal (.32 vs. .43-45).

The interaction Pal × EnvPal shows the lower rate of recognition of the sequence C'C (.33 vs. .42-.44 for all others). This is the only significant difference.

Finally, the three-way interaction Pal × EnvPlace × EnvPal is due to the difference between a sequence of a palatalized coronal plus a plain consonant, t'C, and all the others (pC ($p = .000$), p'C ($p = .001$), tC ($p = .002$), C' ($p = .000$), p'C' ($p = .003$), tC' ($p = .001$), t'C' ($p = .000$). No other differences are found. A plain consonant is less reliably identified when preceded by /t'/ rather than by any other consonants (.24 vs. .40-.48).

In sum, most factors and many interactions are significant. This is partly due to differences between more salient consonants (/p/ and /t'/) and less recoverable ones (/p'/ and /t/) as well as due to the influence of /t'/ on the perception of the following consonant.

6.2.7.2 Discussion

Recall that in our study of the clean stimuli all onset consonants in clusters were well recognized. The only factor that lowered identification was a preceding palatalized segment, particularly in clusters /p't/ and /t't/.

As we saw in the case with single consonants, the identification of onset segments in clusters with noise is much lower than their recognition under ideal conditions. Most of the significant main effects and interactions found in the noisy stimuli are the same as in clean stimuli. The effect, however, is larger for some consonants than others. There is a clear division between more and less salient plain and palatalized consonants. As with single consonants in noise, only two contrasts seem to be distinguished here: plain, associated with /p/, and palatalized, associated with /t'/. The other consonants are lumped together with their counterparts in terms of secondary articulation.

As with the clean onset clusters, we notice the influence of the preceding palatalized consonant on the following plain stop. While in the clean case it was primarily the palatalized labial /p'/ (and then /t'/) that seriously affected the recognition of the following /t/, now it is the palatalized coronal /t'/ that induces the most errors in the perception of /t/. Apparently, the relative robustness of /t'/ is a factor here again. In this case, however, the recovered palatalized gesture/feature is attributed to the following, less salient consonant. In most of the cases /t/ preceded by a palatalized consonant is perceived as a palatalized /t'/ (t'_: 51 tokens; p'_: 36 tokens).

The confusion matrices show that in terms of secondary articulation /t'/ and /p'/ are almost always perceived as palatalized; /t/ is treated also either mainly plain (after plain consonants) or more like palatalized (after the palatalized ones); and /p/ is almost always perceived as plain. The pattern is similar in terms of their place of articulation: /t'/ and /p'/ are most likely to be identified as coronals; /t/ is mostly perceived as coronal, except for the context after /p/, where it is assumed to be labial; /p/ is always identified as labial, however, after /t'/ it is often confused with a coronal consonant.

It appears that the relative position of Tongue Body (recovered from the acoustic signal) seems to determine both palatalization and place decisions. A higher and more front position of TB tends to be associated with both palatalized and coronal consonants, while its lower position is labeled as both plain and labial.

6.2.8 Clusters: Coda (C_1 in VC_1C_2V)

6.2.8.1 Results

Table 6.24 shows the ANOVA results for the identification of onset consonants in clusters. Figure 6.15 presents the mean values for each of the onset consonants in various clusters.

Table 6.24 Correct identification: Coda consonants in clusters (statistics)

	F	p
Place	**30.5**	**0.000**
Pal	**10.33**	**0.005**
EnvPlace	0.63	0.935
EnvPal	0.007	0.935
Place × Pal	**25.02**	**0.000**
Pal × EnvPlace × EnvPal	**5.72**	**0.027**

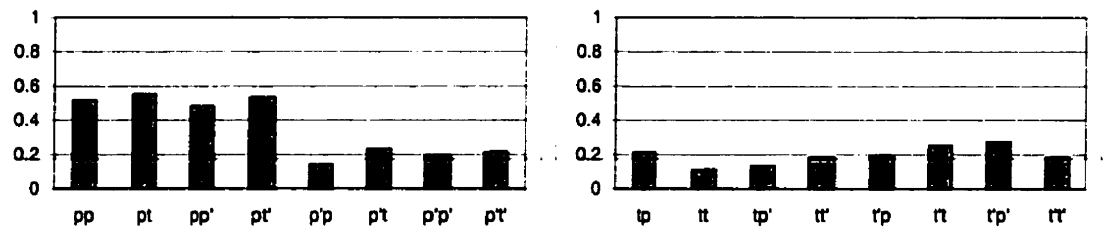


Figure 6.15 Correct identification: Coda consonants (C_1) in clusters

The main effects of Place (the largest effect) and Pal are significant, while EnvPal and EnvPlace are not significant. Labials are more recoverable in noise than coronals (.36 vs. .19) (Place: $p < .001$). Plain consonants are more perceptually robust in noise in this environment than palatalized consonants (.34 vs. .21) (Pal: $p < .01$). These two interactions are mainly due to the difference between /p/ and all the other consonants. Also, identification is different before plain and palatalized consonants (yet for both the rate is the same: .27) (EnvPal: ($p < .001$)).

Two interactions are significant. Place × Pal (the Consonant factor) indicates perception of consonants is different. The post-hoc test determined that only /p/ is significantly different from all other consonants ($p = .000$). It is much better recognized than any other consonant (.52 vs. .16-.22). The three-way interaction (Pal × EnvPlace × EnvPal) is mostly due to the difference between the sequence C'p (i.e. /p'p/ and /t'p/) and all other clusters with C_1 as a plain consonant.[55] It suggests that perception of a palatalized consonant is worse before a plain labial than the recognition of the plain consonant in the same or other environments. It should be noted, however, that the rate for all palatalized consonants in coda is below chance.

In sum, there are a number of significant interactions due to asymmetries in perception of the four consonants in various environments.

6.2.8.2 Discussion

Recall that in the study of clean stimuli we found a much lower overall identification rate in this environment. Recognition was dependent on the quality of the segment: /p/ showed a higher rate of identification, while /p'/ was the worst. It also depended on the secondary articulation of the following consonant: plain consonants were preferred before plain consonants and palatalized before palatalized segments. The place of articulation of C_2 also mattered.

As with the clean stimuli, the results under the current condition show a significant interaction of Place and Pal, pointing to asymmetries in the recognition of consonants. Recall that in the previous case, /p/ in coda had the highest identification, followed by /t'/, and then by /t/ and /p'/. When white noise is added to the stimuli, only /p/ is reliably identified, while all other consonants, including /t'/, are recognized at or below chance. This is similar to the results obtained for single consonants in coda. According to the confusion matrices, these less salient consonants (/p'/, /t/, and /t'/) are most often (but not always) confused with the plain labial /p/. This is a dramatic deterioration of recognition: listeners, who previously could distinguish up to four contrasts in the coda under the ideal conditions, now perceive only one contrast in this position. The default percept is the plain labial, the prototypical non-palatalized segment. Also compare the identification of noisy stimuli in coda to the perception of onset consonants under the same conditions, where two consonants were fairly reliably distinguished.

In sum, the results provide crucial evidence for the preference for a plain rather than palatalized consonant in the preconsonantal coda. This environment is the worst for the plain-palatalized contrast.

6.2.9 Single vs. Clusters

In this section I compare the similarity of identification patterns and RT between single consonants and consonants in clusters combined in terms of syllable position.

The ANOVA results for identification of onset consonants showed that the main effect of Type [$F = 4.941$; $p = .046$] is significant.[56] The identification rate for single onset consonants was higher than for the onset consonants in clusters (.51 vs. 41).

There is no significant effect of Type in the coda position [$F = 2.612$; $p = .132$]. However, there are significant interactions of Place × Type [$F = 18.337$; $p = .001$] and Pal × Type [$F = 9.489$; $p = .010$].[57] The first interaction points to the place asymmetry: labials are better identified when single, while coronals have a higher rate in clusters. It should be noted that the identification rate for both kinds of coronals is below the chance level. The second interaction suggests that single plain consonants are better recognized than the same segments in clusters. It

[55] The sequence C'p is significantly different from Cp ($p = .002$), Cp' ($p = .05$), Ct ($p = .014$), and Ct' ($p = .004$)). The other differences involve C't' vs. Cp ($p = .012$) and Ct' ($p = .019$).

[56] Place is the other significant factor [$F = 4.909$; $p = .047$]. There is also a strong interaction of Place × Pal [$F = 101.931$; $p = .000$], which means that for both types of environments there is a difference in perception of particular consonants: higher recognition of /p/ and /t'/ and lower identification of /p'/ and /t/.

[57] Other significant factors and interactions: Place [$F = 135.277$; $p = .000$], Pal [$F = 78.400$; $p = .000$], and Place × Pal [$F = 96.730$; $p = .000$].

is the reverse for the palatalized consonants. Again, however, the recognition of the former is below the chance level.

This shows that recognition of onset consonants in noise is somewhat dependent on the presence or absence of the preceding consonant. This is not a factor for the coda consonants (single and in clusters). This finding is similar to the results with the clean stimuli.

6.2.10 General discussion

The results again confirm our expectations that the overall recognition of consonants in noise will be lower than under ideal conditions. They also provide additional evidence for the strong perceptual difference between the onset and coda environments, as well as for the asymmetries in terms of place and articulation.

In this section I review the results by looking first at particular consonants and then at environment in general, comparing the noisy results to those found in perception of clean stimuli.

6.2.10.1 Consonants

Figure 6.16 shows the overall identification rate for the four stops (averaged by environment) for the results with and without noise. Both graphs show lower identification rates for /p'/ and /t/ and higher rates for /p/ and /t'/. The differences between these two groups of consonants are more striking in the noisy stimuli, where the perception of /p'/ and /t/ is below chance level. In addition to that, the noisy condition breaks the tie between /p/ and /t'/ in favour of the former.

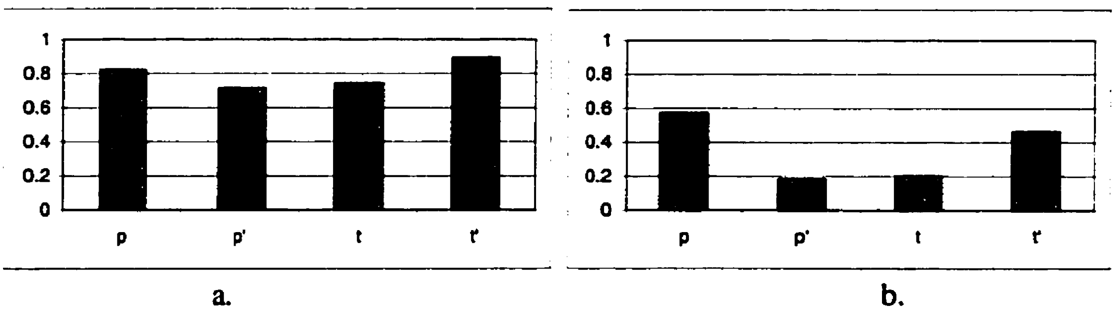
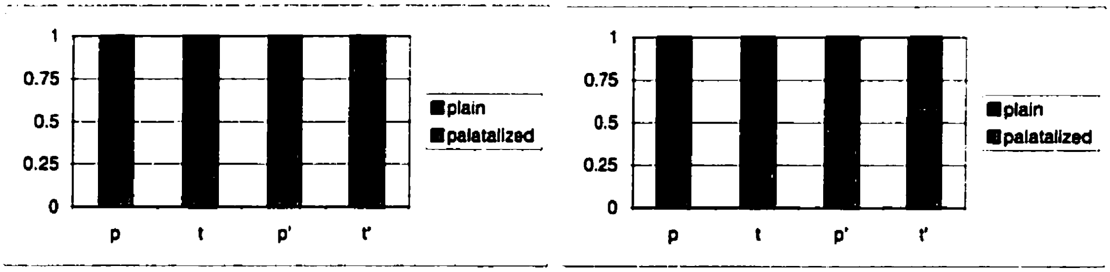


Figure 6.16 Mean consonant identification without noise (a) and with noise (b)

To what degree is each consonant perceived as plain or palatalized in noise (based on Tables 6.18-6.21)? Figure 6.17 shows these consonant palatalization 'profiles' both for the clean (a) and noisy (b) conditions. A very clear distinction between plain and palatalized consonants (a difference of more than .50) in the clean stimuli is to a large degree eroded in noise: the degree of palatalization of palatalized consonants is close to the chance level (.50). The relative difference between the segments is the same under both conditions: /p/ is the most plain and /t'/ is the most palatalized consonant. The other consonants are in between.



a. b.

Figure 6.17 Mean perception of the consonants in terms of palatalization without noise (a) and with noise (b)

These results are averaged for environment. The consonant profiles of plain and palatalized consonants would be more distinct in onset and around the chance level in coda.

6.2.10.2 Environments

Now I turn to the environments to determine the 'best' and the 'worst' contexts for perception of the plain-palatalized contrast in noise and to compare it to the clean condition.

Figure 6.18 shows the correct identification split by correct recognition of the plain-palatalized (palatalization) and labial-coronal (place) contrasts. The environments are ordered from the 'best' for recovering secondary articulation to the 'worst'.

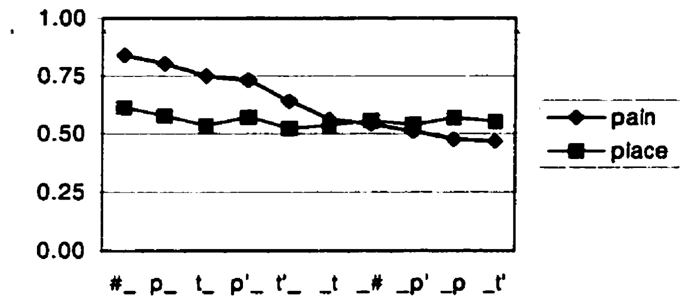


Figure 6.18 Identification of consonants in terms of palatalization (paln) and place in all environments: noisy stimuli

We can see that the palatalization identification rate is quite different from the place identification rate: which does not depend on the environment, hovering slightly above the chance level (.50). The recognition of palatalization, however, is very sensitive to a combination of contextual factors: syllable position, and the nature of the following and preceding consonants in terms of palatalization and place.

Let us compare the recognition of palatalization in noise to the results found with the clean stimuli. In Figure 6.19 the environments are ordered from the most favourable to the least favourable for identification of the clean stimuli.

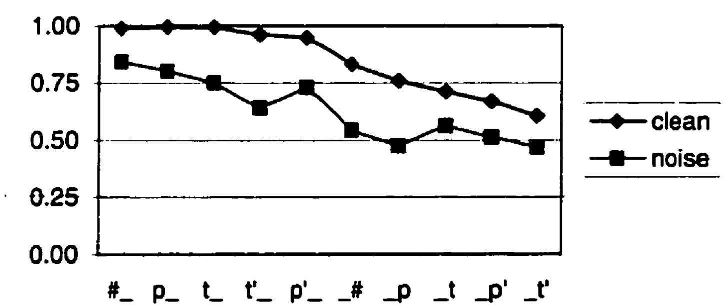


Figure 6.19 Identification of consonants in terms of palatalization in all environments: clean and noisy stimuli

Identification of palatalization in noise shows a pattern very similar to perception under ideal conditions, only at a lower rate. For both we note the influence of syllabic position, preceding and following palatalized consonants. The following differences should be noted. The coda environments in noise show less variability and a much lower identification rate than the onset contexts. The preceding /t'/ affects identification in noise to a greater extent than it normally does. Also, it is the following plain labial rather than coronal that causes more errors in the noisy data.

In sum, the results of the current experiment demonstrate that most of the factors identified under ideal conditions hold true when we add noise to the signal. At the same time, these effects are often more robust in adverse circumstances. The identification in noise sometimes deviates from the clean results (e.g. in identification of place).

The differences between the noisy and clean results are due to masking of some acoustic information available in the signal, particularly stop bursts. As a result of this we find more identification errors and longer processing time.

6.2.7 Conclusion

In this section I investigated the perception by Russian listeners of four stops in a variety of environments under two listening conditions, with and without noise. I found that under both conditions the responses show better identification of the stops in the syllable onset than in the syllable coda. Correct responses are also dependent on the secondary articulation and the place of the following (and, sometimes, preceding) consonant. The stops vary in their perceptual salience and these differences are asymmetrical in terms of palatalization and place. Investigating perception under two conditions was important since it allowed me to determine the extremes of consonant identification. It is reasonable to think that the combination of the results under the two conditions or averaged values of both results would approximate natural human perception, which typically occurs neither in perfect silence, nor in strong noise.

In the next section I investigate the perception of the same clean stimuli by Japanese listeners. This will allow me to determine to what extent (if any) the current results are influenced by language-particular phonotactic knowledge and to what degree they are universal.

6.3 Identification task: Japanese listeners

6.3.0 Introduction

In this section I investigate perception of plain and palatalized stops by non-native speakers, Japanese listeners. The task is phoneme identification.

6.3.1 Background

Japanese has a surface contrast between plain and palatalized consonants. Palatalized consonants are traditionally analyzed as derived from sequences C + j or C + i (Vance 1987: 28-29), sometimes with the exception of coronals (Tsujimura 1996).[58] However, other analyses treat palatalized consonants as having phonemic status (Akamatsu 2000; cf. Trubetzkoy 1958/69: 130, Jakobson 1971). For our purposes I assume that Japanese has a phonemic distinction between plain and palatalized consonants without discussing whether the contrast is derived or not. As we shall see, speakers of the language are well aware of the distinction.

A Japanese consonant inventory is presented in Table 6.25. The contrast between plain and palatalized voiceless stops /p/ vs. /p'/ and /t/ vs. /t'/ will be the focus of our discussion.

Table 6.25 Japanese consonant inventory (based on Akamatsu 1997, 2000; Vance 1987)

Labial		Coronal			Dorsal		Glottal	
plain	pal	plain	pal	pal	plain	pal	plain	pal
p	p'	t	t' [tɕ]		k	k'	[ʔ]	
b	b'	d	d' [dʑ/ʑ]		g	g'		
		ts						
		dz [dz/z]						
ɸ		s	s' [ɕ]				h	h' [ç]
m	m'	n	n' [ɲ]		N [ɰ̃]			
		r [ɹ]	r' [ɹʲ]					
				j	w [ɰ]			

Note: [ʔ] is marginal.

Of particular interest to us are the positional restrictions on the plain-palatalized contrast in Japanese. As in Russian, Japanese plain and palatalized segments are contrastive syllable-

[58] Contrastive palatalization in labials and velars is generally attributed to a subset of lexicon, mimetic vocabulary (e.g. Tsujimura 1996)

initially (Akamatsu 2000).[59] Since the language does not permit final coda obstruents, plain and palatalized consonants do not contrast in final position. In this respect Japanese is very different from Russian. The only obstruent clusters that occur in the language are medial homorganic clusters. In these sequences C_1 agrees with C_2 not only with respect to place, but also with respect to its secondary articulation (Akamatsu 1987). Thus, palatalized consonants are allowed in coda, but restricted only to clusters where the following consonant is a palatalized homorganic segment (i.e. [pʲpʲ] and [tɕtɕ]). Recall that in Russian homorganic clusters also agree with respect to palatalization, while hetero-organic sequences may differ in this respect (e.g. /t'p/, /pt'/, and /tp'/).

These differences between Russian and Japanese are crucial for our study. Investigating the perception of the same Russian stimuli used with native speakers of Russian by Japanese subjects will allow us to answer the question of whether the previous results are due to language-particular influences (frequency or awareness of phonotactic restrictions) or to perceptual properties of the plain-palatalized contrast which are, to a certain extent, language-independent, based on universal recovery of articulatory gestures/features (see section 1.1).

If perception is influenced even partially by language-particular facts, we should expect the responses of Japanese subjects to be different from those of Russian listeners. Particularly, we should expect that there would be no difference between listeners of the different language groups in recognizing /p'/ and /t'/, since both of their corresponding Japanese segments occur initially and in palatalized homorganic clusters, but not finally. The same may be said about all four stops /p/, /p'/, /t/, and /t'/ in initial position: they should exhibit similar behaviour regardless of language. On the other hand, due to the absence of final stops in Japanese we should expect a failure to recognize these segments in the stimuli. In clusters Japanese listeners should be more likely to parse C_1 as identical to C_2 both in place and palatalization.

[59] The contrast, however, is restricted to the environment before back vowels /a/, /o/, and /u/. Only plain consonants

However, if perception of the plain-palatalized contrast is to some extent language-independent, we are likely to find the same asymmetrical patterns with respect to position, place, palatalization, and sensitivity to the place and palatalization of the following or preceding consonant, independent of the language of the listener.

6.3.2 Experimental setup, materials and procedure

Ten native speakers of Japanese[60] were involved in this experiment. None of the listeners had any known hearing disorders and none had prior exposure to the Russian language. All subjects were tested individually. The same stimuli as in the experiment with Russian listeners were used. The listeners were presented only clean stimuli. The response keys were labeled in the *katakana* script, where initial /p/, /p'/ /t/, and /t'/ were presented as *pa, pya, ta, tya*, and the final consonants were transliterated as *pu, pi, to*, and *ti*. This choice of characters was based on a pilot study where Japanese native speakers were asked to write down Russian words with the consonant in question. This was further confirmed with each subject before the session. The procedure was the same as with Russian subjects. Before each session the subjects were given examples of possible stimuli and took a practice trial (about 50 tokens).

The following pronunciation similarities and differences should be noted. The Japanese initial sequences *pa* and *ta* are very similar perceptually to the Russian syllables *pa* and *ta*, disregarding the VOT differences. Based on my observations, in the Japanese syllable *pya* the palatal gesture is timed slightly later than that of Russian /p'a/, however, not as late as in /p'ja/ (see section 6.1.1). The Japanese speakers did not seem to notice the difference in timing. It is interesting that the Japanese speakers unanimously associated Russian initial /t'/ with the Japanese alveolo-palatal affricate [tç] in *tya* or *t(i)*, while Russian speakers tend to categorize the

occur before /e/, and only palatalized ones are found before /i/.

[60] The subjects were 1 male and 9 females. It was not possible have a balanced sample given the time constraints and the overall makeup of the contingent of Japanese University of Toronto students and working holiday visa holders.

Japanese sound differently, closer to Russian palato-alveolar affricate /tʃ/, which is also palatalized phonetically.

Japanese is known for the absence of phonemic final obstruents. In loan words from English, Japanese speakers tend to substitute final stops /p/, /t/, and /k/ with syllables *pu, to,* and *ku* . The same was found in the pilot study with Russian words. In addition, Russian final palatalized consonants /p'/ and /t'/ were associated with the syllables *pi* and *ti*.[61] It is important that the vowels /u/ [ɯ] and /i/ in these syllables tend to be voiceless in Standard Japanese (Vance 1987) in the environments studied here. This makes these sequences virtually phonetically identical to released final stops. The syllable /to/ is different in this respect, since the vowel /o/ is never voiceless.

6.3.3 Analysis

As in the study with Russian listeners, mean identification rates (proportion correct out of 1.00) and response time (RT) were calculated. There was a total of 2000 identification tokens (40 utterances × 5 repetitions × 1 condition × 10 listeners) and the same number of RT tokens. There were 13 cases when subjects did not respond to a stimulus. Due to the high error rate of the coda data, only the RT of onset tokens was examined. As in the study with Russian listeners, I excluded all values lower or higher than two standard deviations from the mean for each subject (based on the overall mean RT values per subject). A total of 28 tokens was excluded. The statistical analysis was the same as applied to the Russian data (see section 6.1.2).

6.3.4 Overall results

In this section I give a general overview of the results. The results by major groups of stimuli are discussed in following sections. Means and standard deviations for correct identifications of

[61] The same method is adopted in transliterating Russian words in Japanese (e.g. in Russian-Japanese dictionaries).

the four stops, based on the average for all Japanese subjects, are presented in Tables 6.26-6.27. The stimuli are divided into groups by syllable position: onset and coda.

Table 6.26 Correct identification: Single consonants (means and standard deviations)

initial	*Mean*	*SD*	*final*	*Mean*	*SD*
ta papy	1.00	0.00	tap apy	0.77	0.19
ta p'apy	0.92	0.14	tap' apy	0.20	0.19
ta tapy	0.92	0.19	tat apy	0.41	0.25
ta t'apy	0.98	0.06	tat' apy	0.70	0.25
mean	*0.96*	*0.10*	mean	*0.52*	*0.22*

Table 6.27 Correct identification: Consonants in clusters (means and standard deviations)

onset	*Mean*	*SD*	*coda*	*Mean*	*SD*
tat papy	1.00	0.00	tap tapy	0.80	0.27
tap papy	0.96	0.13	tap papy	0.70	0.37
tat' papy	0.56	0.22	tap t'apy	0.61	0.30
tap' papy	1.00	0.00	tap p'apy	0.46	0.27
tat p'apy	0.94	0.13	tap' tapy	0.16	0.21
tap p'apy	0.96	0.08	tap' papy	0.16	0.18
tat' p'apy	0.96	0.13	tap' t'apy	0.18	0.26
tap' p'apy	0.98	0.06	tap' p'apy	0.51	0.27
tap tapy	1.00	0.00	tat papy	0.38	0.27
tat tapy	0.96	0.08	tat tapy	0.30	0.22
tap' tapy	0.96	0.08	tat p'apy	0.14	0.16
tat' tapy	0.98	0.06	tat t'apy	0.20	0.23
tap t'apy	0.96	0.08	tat' papy	0.40	0.29
tat t'apy	0.98	0.06	tat' tapy	0.36	0.36
tap' t'apy	0.92	0.14	tat' p'apy	0.26	0.27
tat' t'apy	0.92	0.19	tat' t'apy	0.70	0.25
mean	*0.94*	*0.09*	mean	*0.40*	*0.26*

The total average identification rate was .68 (.17) (compare to .79 (.15) for Russian subjects). The rate for single consonants was 0.74 (0.16) and for consonants in clusters was .67 (.18). Within the single consonants a rate of .96 (.10) was found for the onset (initial) and a rate of .52 (.22) for the coda (final). In clusters the onset consonants had a rate of .94 (.09) and the coda consonants a rate of .40 (.26). Collapsing all the groups into onset and coda classes, the total rate for the onset is .94 (.09), and the total rate for the coda is .42 (.25).

Identification matrices for /p/, /p'/, /t/, and /t'/ are shown in Tables 6.28-6.31.

Table 6.28 Confusion matrix: /p/ (tokens)

		Consonant				Place		Palatalization	
		p	p'	t	t'	lab	cor	npal	pal
single	a._a	50	0	0	0	50	0	50	0
	a_.a	38	2	6	6	40	12	44	8
CC-onset	t_a	50	0	0	0	50	0	50	0
	p_a	48	1	1	0	49	1	49	1
	t'_a	47	0	2	0	47	2	49	0
	p'_a	49	0	1	0	49	1	50	0
CC-coda	a_t	40	1	7	2	41	9	47	3
	a_p	36	7	2	5	43	7	38	12
	a_t'	29	7	3	9	36	12	32	16
	a_p'	23	20	2	9	43	11	25	29
	total	410	38	24	31	448	55	434	69
	ratio	0.82	0.08	0.05	0.06	0.89	0.11	0.86	0.14

Table 6.29 Confusion matrix: /p'/ (tokens)

		Consonant				Place		Palatalization	
		p	p'	t	t'	lab	cor	npal	pal
single	a._a	2	45	1	1	47	2	3	46
	a_.a	27	10	6	10	37	16	33	20
CC-onset	t'_a	2	46	0	2	48	2	2	48
	p'_a	2	46	0	2	48	2	2	48
	t_a	1	48	1	0	49	1	2	48
	p_a	1	49	0	0	50	0	1	49
CC-coda	a_t	28	8	6	8	36	14	34	16
	a_p	33	8	5	3	41	8	38	11
	a_t'	20	9	5	15	29	20	25	24
	a_p'	12	25	5	9	37	14	17	34
	total	128	294	29	50	422	79	157	344
	ratio	0.26	0.59	0.06	0.10	0.84	0.16	0.31	0.69

Table 6.30 Confusion matrix: /t/ (tokens)

		Consonant				Place		Palatalization	
		p	p'	t	t'	lab	cor	npal	pal
single	a._a	2	1	45	1	3	46	47	2
	a_.a	20	2	20	7	22	27	40	9
CC-onset	p_a	0	0	50	0	0	50	50	0
	t_a	2	0	48	0	2	48	50	0
	p'_a	1	2	28	19	3	47	29	21
	t'_a	0	0	50	0	0	50	50	0
CC-coda	a_p	21	5	19	6	26	25	40	11
	a_t	24	4	15	8	28	23	39	12
	a_p'	25	13	7	4	38	11	32	17
	a_t'	10	3	10	23	13	33	20	26
	total	105	30	292	68	135	360	397	98
	ratio	0.21	0.06	0.59	0.14	0.27	0.73	0.80	0.20

Table 6.31 Confusion matrix: /t'/ (tokens)

		Consonant				Place		Palatalization	
		p	p'	t	t'	lab	cor	npal	pal
single	a._a	0	1	0	**49**	1	**49**	0	**50**
	a_.a	5	3	7	**31**	8	**38**	12	**34**
CC-onset	p'_a	0	2	0	**48**	2	**48**	0	**50**
	t'_a	1	0	0	**49**	1	**49**	1	**49**
	p_a	0	3	0	**47**	3	**47**	0	**50**
	t_a	0	2	0	**48**	2	**48**	0	**50**
CC-coda	a_p	**19**	6	5	17	**25**	22	**24**	23
	a_t	16	5	11	**18**	21	**29**	**27**	23
	a_p'	11	**19**	7	13	**30**	20	18	**32**
	a_t'	9	1	5	**34**	10	**39**	14	**35**
	total	*61*	*42*	*35*	***354***	*103*	***389***	*96*	***396***
	ratio	*0.12*	*0.09*	*0.07*	***0.72***	*0.21*	***0.79***	*0.20*	***0.80***

Tables 6.32-6.33 I present the RT results by the three groups of stimuli. The total average RT was 1479 (397) ms (compared to 1464 (237) ms for Russian listeners). The result for single consonants was 1482 (439) ms and for consonants in clusters was 1477 (356). Within the single consonants a rate of 1234 (287) ms was found for the onset (initial) and a rate of 1729 (590) ms for the coda (final). In clusters the onset consonants had a rate of 1370 (259) and the coda segments a rate of 1584 (453) ms. Collapsing all the groups into onset and coda classes, the total rate for the onset is 1302 (273) ms, and the total rate for the coda is 1656 (521) ms.

Table 6.32 Response Time: Single consonants (means and standard deviations)

initial	*Mean*	*SD*	*final*	*Mean*	*SD*
ta **papy**	1208	209	**tap** apy	1653	446
ta **p'apy**	1171	273	**tap'** apy	1588	561
ta **tapy**	1224	165	**tat** apy	2004	529
ta **t'apy**	1333	226	**tat'** apy	1857	509
	1234	*218*		*1776*	*511*

Table 6.33 Response Time: Consonants in clusters (means and standard deviations)

onset	*Mean*	*SD*	*coda*	*Mean*	*SD*
tat papy	1256	230	tap tapy	1400	449
tap papy	1336	261	tap papy	1542	446
tat' papy	1298	212	tap t'apy	1536	619
tap' papy	1332	145	tap p'apy	1414	392
tat p'apy	1290	256	tap' tapy	1542	667
tap p'apy	1271	138	tap' papy	1300	708
tat' p'apy	1436	217	tap' t'apy	1562	441
tap' p'apy	1336	255	tap' p'apy	1714	441
tap tapy	1305	213	tat papy	1552	408
tat tapy	1770	432	tat tapy	1688	461
tap' tapy	1285	182	tat p'apy	1513	519
tat' tapy	1414	157	tat t'apy	2067	358
tap t'apy	1423	254	tat' papy	1738	631
tat t'apy	1422	284	tat' tapy	1790	363
tap' t'apy	1435	249	tat' p'apy	1773	665
tat' t'apy	1307	133	tat' t'apy	1443	234
	1370	*226*		*1598*	*488*

In the next sections I present the results in more detail.

6.3.5 Single consonants

6.3.5.1 Results

Table 6.34 shows the ANOVA results for the identification of onset consonants in clusters. Figure 6.20 presents the mean values for each of the onset consonants in various clusters.

Table 6.34 Correct identification: Single consonants (statistics)

	F	p
Place	0.55	0.475
Pal	3.35	0.101
Env	**117.25**	**0.000**
Place × Pal	**80.36**	**0.000**
Pal × Env	**6.88**	**0.027**
Place × Pal × Env	**24.1**	**0.001**

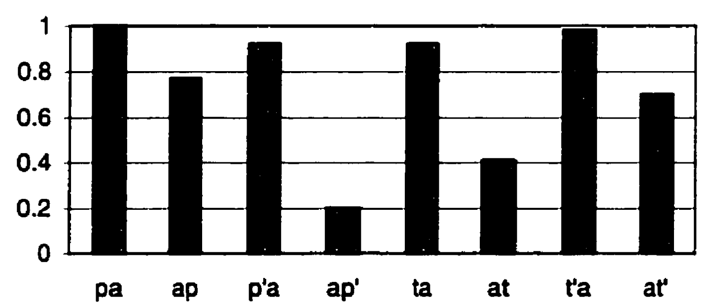


Figure 6.20 Correct identification: Single consonants

Of all the main effects, Env is significant, while Place and Pal are not significant. Correct identification of consonants in coda (.52) was significantly lower than in onset (.96) (Env: [$p < .001$]).

All two- and three-way interactions of the factors were also significant. The interaction Place × Pal is of particular interest to us. A Tukey HSD post-hoc test determined that perception of plain and palatalized consonants (both labials and coronals) is significantly different. However, this relation within a particular place of articulation is different. Plain labials have a higher rate than the palatalized ones (.89 vs. .56) ($p = .000$). On the other hand, the palatalized coronal /t'/ is more easily recovered than the plain /t/ (.84 vs. .67) ($p = .007$). Also, /p/ is different from /t/ ($p = .002$) and /t'/ is different from /p'/ ($p = .000$). There is no difference in identification between /p/ and /t/. The interaction Pal × Env shows that consonants are significantly less recognizable in the coda than in the onset, irrespective of their secondary articulation (both $p = .000$). At the same time, the coda palatalized consonants are less recoverable than the coda plain segments ($p = .014$). The rate of identification of these segment types is as follows: #C (0.96), C# (0.59), #C' (0.95), and C'# (0.45).

Finally, the interaction Place × Pal × Env suggests an inconsistent recognition of segments depending on syllabic position. The analysis shows that this is due mostly to the differences of final consonants (especially /t/ and /p'/) from the same consonants in the initial position and some other consonants. The final /p'/ and /t/ have the lowest identification rate (.2 and .41) and the initial /p/ and /t'/ are the best recognized (1.00 and .98).

In sum, Env is the most important factor. The significant interactions are due to the different degree of recognition of the final consonants, particularly, lower identification of /p'/.

Table 6.35 shows the ANOVA results of the reaction time (RT) for onset consonants in clusters. Here and further I present only significant interactions. Figure 6.21 provides the mean values for each of the onset consonants in various clusters.

Table 6.35 RT: Single consonants (statistics)

	F	p
Place	**10.64**	**0.022**
Pal	.23	0.647
Env	**18.25**	**0.008**

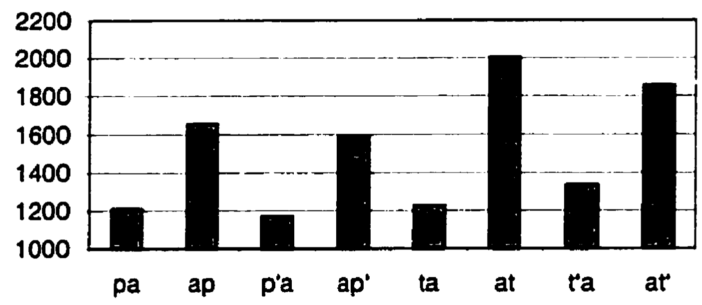


Figure 6.21 RT (ms): Single consonants

The main effects of Env and Place are significant. It takes more than 500 ms longer for the Japanese listeners to correctly identify a final consonant compared to an initial consonant (1234 vs. 1776 ms) (Env: ($p < 0.01$)). It also takes more time to recognize coronals than labials (1405 vs. 1605 ms) (Place: ($p < 0.05$)). The factor of Pal is not significant.

The factor that is significant for both identification and RT is Env.

6.3.5.2 Discussion

Recall that in the experiments with Russian listeners all consonants in onset showed very high identification rates. Perception of the same consonants in coda was lower, particularly of /p'/. Adding noise lowered overall identification: only two consonants were reliably identified in the onset (/p/, /t'/) and only one (/p/) in the coda.

The results obtained from the Japanese listeners show similarities with both clean and noisy results for Russian listeners. The accuracy of perception of the same consonants in onset and in coda are very different. The identification of onset consonants is as high as in the Russian clean data. The rate in the coda is low; it is lower than in the clean results, however, it is higher than the identification in noise by Russian subjects. The consonants that are affected most in the coda are /p'/ and /t/, the same segments whose recognition is most damaged in noise (Russian listeners). /p'/ has the lowest identification rate, much as for Russian listeners under both conditions. When misidentified, both /p'/ and /t/ are most often confused with the plain labial /p/. Recall that the same confusion pattern was found in the Russian results in noise (and in clean stimuli for /p'/).

The RT difference between onset and coda consonants, found for Russian listeners, is even more substantial for Japanese subjects. Processing of place of articulation, however, is different. While for Russians labials (especially /p/) take longer to process, it is coronals that take more time to recognize for the Japanese listeners. In the case of /t'/ this may be explained by phonetic differences between the Russian palatalized coronal and the Japanese /t'/ [cç] (see section 6.3.1). As for [t], recall that Japanese uses the syllable *to* to represent this sound finally. Since the vowel [o] in *to* is never voiceless, the Russian [t] without a release is more problematic for the Japanese listeners. This may have also contributed to the lower identification rate for the final /t/ for Japanese, but not for Russian subjects.

To summarize, the overall results for Japanese listeners are in the same direction as those for Russian subjects, and show that perception of plain and palatalized consonants by Japanese listeners is very similar to that of Russian subjects.

6.3.6 Clusters: Onset (C_2 in VC_1C_2V)

6.3.6.1 Results

Table 6.36 shows the ANOVA results for the identification of onset consonants in clusters. Figure 6.22 presents the mean values for each of the onset consonants in various clusters.

Table 6.36 Correct identification: Onset consonants in clusters (statistics)

	F	p
Place	1.6	0.238
Pal	1.5	0.252
EnvPlace	**24.75**	**0.001**
EnvPal	**61.71**	**0.000**
Place × Pal	**11.59**	**0.008**
Place × EnvPlace	**8.78**	**0.016**
Pal × EnvPlace	**9.47**	**0.013**
Place × EnvPal	**6.61**	**0.030**
Pal × EnvPal	**21.44**	**0.001**
EnvPlace × EnvPal	**17.02**	**0.003**
Place × Pal × EnvPlace	**11.29**	**0.008**
Place × Pal × EnvPal	**23.05**	**0.001**
Place × EnvPlace × EnvPal	**11.18**	**0.009**
Pal × EnvPlace × EnvPal	**12**	**0.007**
Place × Pal × EnvPlace × EnvPal	**17.82**	**0.002**

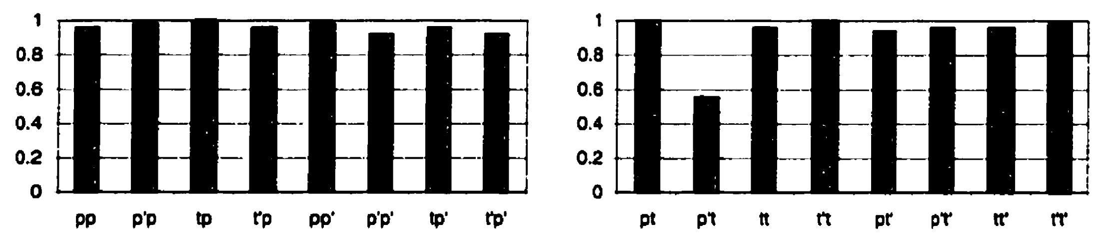


Figure 6.22 Correct identification: Onset consonants (C_2) in clusters

The main effects of EnvPal and EnvPlace are significant. Consonants in the onset position are better identified when preceded by plain rather than palatalized consonants (.97 vs. .91) (EnvPal: $p < .001$). They are more easily recognized if the preceding consonant is coronal (.97 vs. .92) (EnvPlace: $p < .01$).

All factor interactions were significant. Here I discuss only the four-way interaction (Place × Pal × EnvPlace × EnvPal), since all other interactions seem to follow from it. This interaction suggests that perception of a particular consonant is different depending on the place and secondary articulation of the preceding consonant. The post-hoc test showed that this interaction is due entirely to the contrast of /p't/ with all other clusters ($p = .000$ or $.001$). This sequence shows the lowest rate, .56, while all other sequences are in the range .92 to 1.00.

In sum, all the factors and all interactions of these factors are significant. Most of the interactions are due to the lower recognition of /p'/ in the coda, as well as /p/ and /t'/ in this environment.

Table 6.37 shows the ANOVA results of the reaction time (RT) for onset consonants in clusters. Here and further I present only significant factor interactions. Figure 6.23 provides the mean values for each of the onset consonants in various clusters.

Table 6.37 RT: Onset consonants in clusters (statistics)

	F	p
Place	**8.01**	**0.020**
Pal	.06	0.81
EnvPlace	**8.74**	**0.016**
EnvPal	**12.10**	**0.007**
Pal x EnvPlace	**8.66**	**0.016**
Pal x EnvPal	**11.43**	**0.008**
Place x Pal x EnvPal	**31.73**	**0.000**
Place x EnvPlace x EnvPal	**5.90**	**0.038**

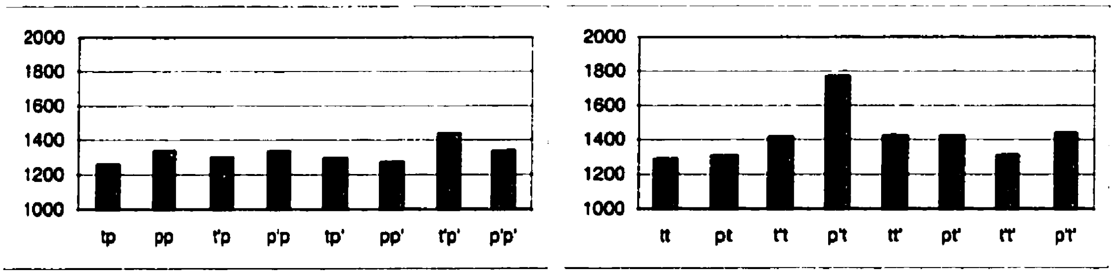


Figure 6.23 RT (ms): Onset consonants (C_2) in clusters

The main effects of EnvPal, EnvPlace, and Place are significant. Consonants in this position are identified faster when preceded by plain rather than palatalized consonants (1323 vs. 1416 ms) (EnvPal: $p < .01$). They are more easily recognized if the preceding consonant is coronal (1338 vs. 1401 ms) (EnvPlace: $p < .05$). It takes less time to correctly identify a labial than a coronal (1319 vs. 1420 ms) (Place: $p < .05$).

All of the interactions are significant. The three-way interaction of Place x Pal x EnvPal suggests that RT for a consonant varies depending on the secondary articulation of the preceding consonant. The RT for a /t/ preceded by a palatalized consonant, C't, is significantly higher than for a /t/ preceded by a plain segment, Ct (1592 vs. 1295 ms) ($p = .001$). In fact, the sequence C't is different from all other combinations. No other differences are found.

All the other interactions seem to be the result of the difference between the /p't/ cluster and all the other clusters (Pal × EnvPal: C'C vs. CC ($p = .001$), CC' ($p = .021$); Pal × EnvPlace: labial-C vs. coronal-C only .010; Place × EnvPlace × EnvPal: p'-coronal vs. all) (p-coronal .044, t-coronal .035, t'-coronal .041).

The factors and interactions that are significant for both identification and RT are as follows: EnvPlace, EnvPal, Pal × EnvPlace, Pal × EnvPal, Place × Pal × EnvPal, Place × EnvPlace × EnvPal.

6.3.6.2 Discussion

Recall that in the experiment with Russian listeners without noise all onset consonants in clusters were well recognized. The only factor that lowered identification was a preceding palatalized segment, in clusters /p't/ and /t't/. Adding noise led to confusion of consonants in terms of their place of articulation (/p'/ with /t'/ and /t/ with /p/). The effect of the preceding palatalized consonant was also significant.

As with single consonants, the current results with Japanese listeners show the pattern found in the clean Russian stimuli: very high correct identification (sometimes better than that of native speakers) of onset consonants. The only exception is the low recognition of /t/ after a palatalized /p'/, an effect also found in the Russian results. The difference is in the magnitude of the error: while /t/ in this cluster is confused by Russian subjects 18% of the time, the Japanese listeners mishear it in 44% of the cases. In most of the confused responses the consonant was identified as the palatalized /t'/ by both Russians and Japanese. This cluster is also the worst in terms of RT results for both groups of listeners.

It is interesting, however, that unlike in the Russian identification and RT results, in the Japanese results the preceding /t'/ does not have any effect on the following homorganic /t/. On the contrary, /t/ in this cluster has the highest recognition rate of 1.00 and an average RT of 1414 (157) ms.

In general, the Japanese and the Russian speakers show a very similar identification of onset consonants in clusters. This further confirms our expectations that the fundamentals of perception, as recovery of gestures/features, are universal, with language-particular properties entering at higher levels.

6.3.7 Clusters: Coda (C_1 in VC_1C_2V)

6.3.7.1 Results

Table 6.38 shows the ANOVA results for the identification of onset consonants in clusters. Figure 6.24 presents the mean values for each of the onset consonants in various clusters.

Table 6.38 Correct identification: Coda consonants in clusters (statistics)

	F	p
Place	4.3	0.068
Pal	1.88	0.203
EnvPlace	1.18	0.305
EnvPal	0.5	0.499
Place × Pal	**39.55**	**0.000**
Pal × EnvPal	**11.33**	**0.008**
EnvPlace × EnvPal	**12.63**	**0.006**
Place × Pal × EnvPlace	**22.61**	**0.001**
Place × EnvPlace × EnvPal	**15.58**	**0.003**

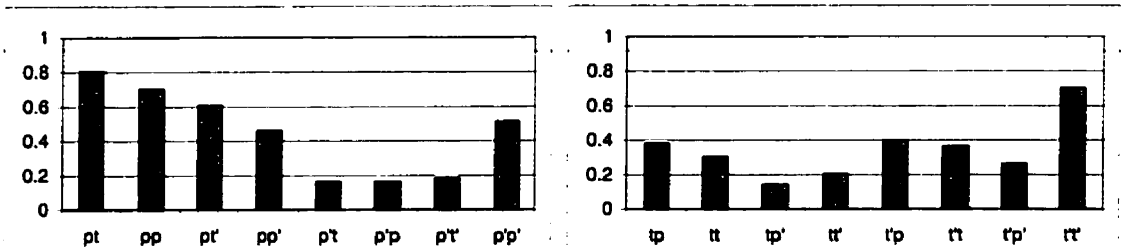


Figure 6.24 Correct identification: Coda consonants (C_1) in clusters

None of the main effects (Place, Pal, EnvPlace, and EnvPal) are significant. However, most of the interactions show significant effects. The major effect is that of Place × Pal (the Consonant factor). The post-hoc test analysis shows that the difference is only between /p/ and all other consonants (/p/ vs. /p'/ (p = .001), /t/ (p = .001), /t'/ (p = .036)). /p/ shows the highest identification rate: .64, while the other consonants are substantially less reliably identified (/p'/ (0.25), /t/ (0.26), /t'/ (0.43)).

The analysis of the interaction Pal × EnvPal shows that perception of the sequence of two plain consonants (CC: .55) is significantly higher than identification of the palatalized consonant followed by a plain one (C'C: .27) (p = .016). Other differences are not significant.

Other interactions show dependence of perception results on the quality of the following consonant (/p'/ vs. all other consonants) and on the differences between both C_1 and C_2. There are no other significant differences between the same coda segments in various contexts (e.g. /t'p/ vs. /t't/ or /t't'/), except for the sequences coronal-p' vs. coronal-t' (p = .024).

The mean value for RT in coda is 1598 (488). The details were not investigated, since the identification rate in this environment was relatively low (.40).

In sum, a number of interactions of factors are significant. They point to asymmetries in the perception of particular coda consonants and dependence of it on the quality of the following consonant.

6.3.7.2 Discussion

Recall that in the experiment with Russian listeners we found a much lower overall identification in the coda environment. Recognition was dependent on the quality of the segment: /p/ showed a higher rate, while /p'/ was the worst. It also depended on the secondary articulation of the following consonant: plain consonants were preferred before other plain consonants and palatalized before other palatalized segments. The place of articulation of C_2 also was significant. Adding noise led to overall poorer recognition and a perceptual bias toward the plain labial /p/.

The Japanese results for coda position in clusters are somewhere between the clean and noisy results for the Russian subjects. The overall rate of identification is lower than for native speakers performing under ideal circumstances, while it is higher than in their performance in noise. As before, we find the same relative salience of consonants in coda: /p/ is the best and /p'/ is the worst. Most of the time the latter is confused with the plain labial. The same is also true for /t/ and, to some extent, for /t'/ (which is perceived as /p/ before another /p/ and as /t/ before a /t/). This is very close to the pattern found in Russian perception in noise.

As in the case with the Russian listeners (clean stimuli), the Japanese are sensitive to the secondary articulation of the following consonants, showing a preference for the sequences that agree in terms of palatalization. They also show the effect of place of articulation of the following palatalized consonant. Thus, /p/ and /t/ are most likely to be confused with their palatalized counterparts when they are followed by /p'/ and /t'/, particularly if the following segments are homorganic. However, the effect of place of the following plain consonant on palatalization of C_1 does not seem to influence the Japanese listeners. This is contrary to the clean results, while consistent with the noisy results for Russian subjects.

6.3.8 Single vs. Clusters

The ANOVA results for identification of onset consonants showed that there was no significant difference between single and cluster consonants in the onset position (the main effect of Type) [F = .062; p = .808].[62] In the coda environment there is no difference either in terms of Type [F = 2.066; p = .176][63], or in terms of significant interactions with Type.

The ANOVA results for RT in onset also suggest the similarity between the two types of consonants: Type is not significant [F = 3.861; p = .073].[64] Interactions with type are not significant either. The RT in coda was not investigated due to scarce data.

This shows that the major factor in perception of stops by the Japanese listeners is syllable position. This is similar to the results found for Russian subjects.

[62] Neither the main effects of Place [F = .169; p = .688] and Pal [F = .016; p = .901], nor any interactions with them, were significant.

[63] Place [F = .040; p = .845] and Pal [F = 2.005; p = .182] are not significant. The interaction Place × Pal is significant [F = 16.557; p = .002]. It shows the similar relations between plain and palatalized labials and coronals: better identification of /p/ and /t'/ (especially /p/) and higher error rates for /p'/ and /t/.

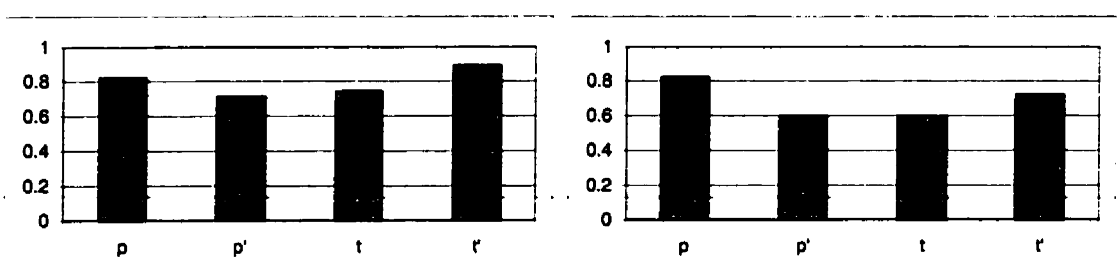
6.3.9 General discussion

The overall results confirm the hypothesis that recognition of consonants by non-native listeners is not entirely different from the performance of the Russian subjects. The Japanese listeners exhibit a somewhat lower identification rate, mostly in coda, however, the factors that affect perception in a variety of environments are almost the same. The same perceptual differences between onset and coda environments and asymmetries in terms of place and articulation noticed with the Russian listeners are found with the Japanese listeners as well.

In this section I review the results by looking first at particular consonants and then at environment in general. I compare the current results to the ones found for the Russian subjects.

6.3.9.1 Consonants

Figure 6.25 shows the overall identification rate for the four stops averaged by environment for the results for Russian (clean) and Japanese listeners. Both graphs show lower identification rates for /p'/ and /t/ and higher recognition of /p/ and /t'/. /p/ is the best identified segment for the Japanese subjects, which is also the case in the Russian noisy results. As expected, the overall rate for non-native speakers is lower than for the Russians with the exception of /p/.



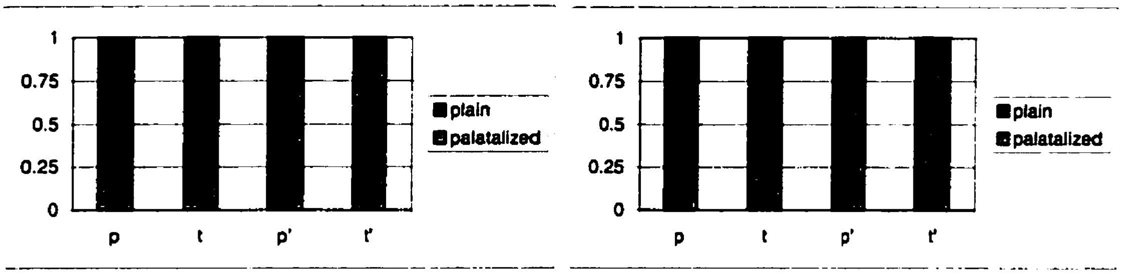
a. b.

Figure 6.25 Mean consonant identification: Russian (clean) (a) and Japanese (b) listeners

Let us compare the degree of palatalization of the stops for both groups of listeners, i.e. to what degree each consonant is perceived as plain or palatalized in noise (based on Tables 6.18-6.21). Figure 6.26 shows the results for both the Russians (a) and the Japanese (b). For both

[64] None of the other factors or interactions is significant either for Place [F = 1.900; p = .193], or Pal [F = .038; p = .848]).

groups we find a clear distinction between plain and palatalized consonants, although the palatalized consonants are more often perceived as plain by the Japanese subjects. The relative difference between the segments is the same for both groups: /p/ is more plain than /t/ and /t'/ is more palatalized than /p'/.



a. b.

Figure 6.26 Mean perception of the consonants in terms of palatalization: Russian (clean) (a) and Japanese (b) listeners

These results are averaged across environment. Recall that the Japanese show almost perfect identification in onset, but worse than native recognition in coda. In other words, plain consonants are 'more plain' and palatalized segments are 'more palatalized' in onset, but their distinctness is blurred in the coda.

6.3.9.2 Environments

Do both Japanese and Russian listeners show the same most and least favourable contexts for perception of the plain-palatalized contrast?

Figure 6.27 plots the correct identification by Japanese subjects split by correct recognition of the plain-palatalized and labial-coronal contrasts. The environments are ordered from the 'best' for recovering secondary articulation to the 'worst'.

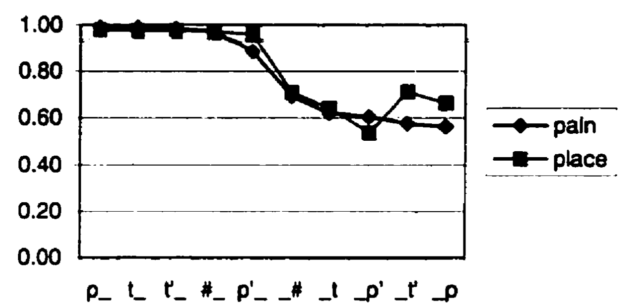


Figure 6.27 Identification of consonants in terms of
Palatalization (paln) and place in all environments: Japanese listeners

It can be seen that correct identification of palatalization is very similar to the recognition of place: they both are almost perfect in onset and are fairly low in coda. The only major differences between the two are before _t' and _p, where place is better identified.

Below (Figure 6.28) I compare the recognition of palatalization by Japanese subjects (referred to as JCI: Japanese clean identification) to the perception of it by Russian subjects under two conditions, clean (RCI: Russian clean identification) and noisy (RNI: Russian noisy identification) stimuli. The environments are ordered from the 'best' to the 'worst' for the Russian clean stimuli.

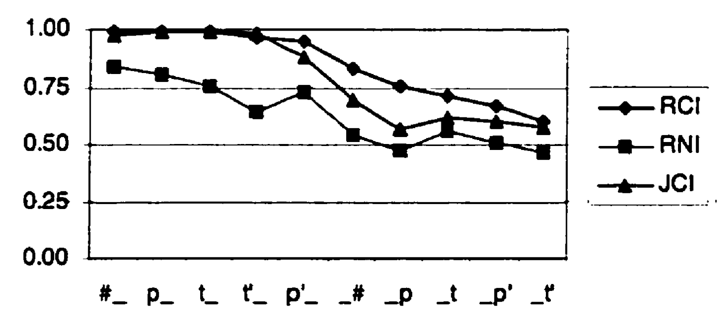


Figure 6.28 Identification of consonants in terms of palatalization and place
in all environments: Russian clean (RCI) and noisy (RNI) identification and Japanese clean identification (JCI)

The perception of palatalization by Japanese subjects is almost identical to that of Russian listeners without noise (the only difference is the stronger effect of /p'/ in the following consonant for the Japanese). The rate of identification of palatalization by the Japanese in coda is lower than in the clean Russian results, but higher than in noise. In fact, it follows the noise results very closely.

This comparison of results obtained under three different conditions provides important evidence for the fact that perception of palatalization is very similar for both Russian and Japanese listeners. The noise for Russian speakers plays the same role in identification as less familiarity with non-native phonotactic patterns for the Japanese subjects.

6.3.10 Conclusion

The results of the perception study of the clean stimuli by Japanese listeners allowed us to determine whether they show the same effects as the Russian subjects. The findings reveal that the two groups of listeners behave very similarly. However, the similarities are more striking if one takes into account the two conditions. The responses of the Japanese subjects in onset position match the Russian clean results, but they tend to agree with the Russian noisy responses in the coda. These findings provide evidence for the fact that the main factors in perception of palatalization found across the two groups of listeners are to a large degree language-independent and motivated by general limitations on the recoverability of gestures/features.

6.4 Writing task: Russian listeners

6.4.0 Introduction: Expectations

The experiments discussed in sections 6.1-6.3 had certain methodological limitations. Particularly, the listeners were limited in their choice of segments (only /p/, /p'/, /t/, and /t'/) and they focused on only two positions in the utterance, medial coda and onset consonants. The current experiment was designed to determine whether these four utterance-medial stops can exhibit a wider range of perceptual variation, and whether their quality affects other segments in the utterance. The task of writing down the stimuli in orthography (the writing task) may provide answers these questions, since it does not impose strict limitations on listeners (other than the choice of phonemes and language orthography). At the same time the responses are more likely to depend on language-specific knowledge of phoneme frequency and phonotactics. In addition, orthography is rarely perfect in encoding the relevant phonotactic nuances. Overall, the use of both methods is likely to provide a more accurate view of perception of consonants.

6.4.1 Experimental setup

6.4.1 Materials and procedure

The same Russian subjects who participated in the experiments discussed in sections 6.1-6.3 were involved in this experiment. The test utterances were same clean stimuli (*ta*C_1*#* C_2*apy*), randomized and mixed with filler words. Due to time limitations each utterance was repeated only twice (two different tokens of the same utterance). The stimuli were presented using PsyScope with an inter-stimulus interval of 5000 ms. Listeners were instructed that they would be presented with 2-word nonsense utterances and their task was to write down these utterances in Russian orthography on answer sheets. No examples of stimuli were given prior to the sessions.

6.4.2 Analysis

A total of 160 written tokens were collected (40 utterances × 2 repetitions × 20 listeners) for the Russian subjects. The responses were transcribed and entered on a spreadsheet. The results were not analyzed statistically due to the relatively small number of tokens and the descriptive nature of the task. Note that certain perceptual differences discussed below may not always be statistically significant.

6.4.3 Results

I address two questions in this section: how are consonants and vowels of the stimuli utterances perceived in general, and how are they identified in particular environments? These questions divide the section into two parts: the general perception of segments (6.4.3.1) and their perception in terms of environments (6.4.3.2).

6.4.3.1 Segments

In this sub-section I discuss the range of variation in perception of consonants and vowels of the presented stimuli. All tokens can be seen as a sequence $C_1V_1C_2C_3V_2C_4V_3$, where C_1 is /t/, V_1 and V_2 are /a/, C_4 is /p/, V_3 is /i/ [i], and C_2 and C_3 are represented by any of the four consonants: /p/, /p'/, /t/, and /t'/. Since only the first four segments show relevant variation in the perception data, our discussion is limited to them. I begin by presenting the results for C_2 and C_3, followed by C_1 and C_4, and V_1 and V_2.

6.4.3.1.1 C_2 and C_3

Table 6.39 presents the confusion results for the coda consonant C_2 and the onset segment C_3.

Table 6.39 Confusion matrix: C_2 and C_3 as segments (tokens)

		plain								palatalized									Ø
		lab		cor					vel	lab	cor					jC/Cj			
	input	p	f	t	d	ts	s	z	k	p'	t'	d'	ts'	tʃ	j	jp	jt	t'j	
C_2	p	**120**	1	10	3				2	12	8				4				20
	p'	**92**		8					5	25	7				21	10	1		11
	t	22		**66**	5	2	3		3	2	35	2			6	2	1		31
	t'	8		32	2	3		1	1	2	**66**	4	1	1	24	8	4		23
	total	*242*	*1*	*116*	*10*	*5*	*3*	*1*	*11*	*41*	*116*	*6*	*1*	*1*	*55*	*20*	*6*	*0*	*85*
C3	p	**160**		2						1	2								15
	p'	15		5						**145**	6				2				6
	t	9		**130**		6				2	15				1				17
	t'	1		4		1				1	**160**		1		7			2	3
	total	*185*	*0*	*141*	*0*	*7*	*0*	*0*	*0*	*149*	*183*	*0*	*1*	*0*	*10*	*0*	*0*	*2*	*41*

The input segments were /p/, /p'/, /t/, and /t'/. They are given in the leftmost column. These consonants are perceived by listeners in a variety of ways: as segments, sequences of segments, or missed altogether (Ø; the rightmost column). These are organized in terms of secondary articulation and place. Sequences of consonants with yod are grouped together with palatalized consonants. The numbers represent the total of tokens, or responses, for a particular segment regardless of the environment. The maximum numbers for each input segment are given in bold. For example, when presented with utterances with a consonant /p/ in the coda position, Russian listeners correctly identified it as /p/ in 120 cases; in other instances (60 cases) they heard it as /p'/ (12), /t/ (10), /t'/ (8), /j/ (4), /k/ (2), etc. or did not notice the consonant at all (20). At the same time, the same consonant /p/ in the onset was recognized 160 times and misidentified or missed in only 20 cases (the rightmost column).

Overall, the four stops, /p/, /p'/, /t/, and /t'/ are perceived as 14 segments (and 3 segment combinations) out of 35 possible consonants (Table 6.40). The choice of the consonants is not random: the responses are limited to voiceless and voiced obstruents (stops, affricates, and fricatives) and the glide /j/. None of the listeners identified the stops as nasals (/m/, /m'/, /n/, and /n'/), laterals (/l/, /l'/) or rhotics (/r/, /r'/). The only sonorant found in the responses is /j/. Confusion in terms of voice among obstruents is limited to coronals only.

Table 6.40 Russian consonant inventory and perception of the stops /p/, /t/, /p'/, /t'/; the shaded areas represent listeners' responses

labial		coronal				dorsal	
plain	pal	plain	pal	plain	pal	plain	pal
p	p'	[illegible]	[illegible]			[illegible]	k'
b	b'	d	d'			g	g'
		[illegible]			[illegible]		
[illegible]	f'	[illegible]	s'	ʂ	ʃ:	x	[x']
v	v'	[illegible]	z'	ʐ			
m	m'	n	n'				
		l	l'				
		r	r'				
					[illegible]		

In order to analyze these data I combine the perceived consonants in several ways. All the responses can be divided into three groups (Table 6.41) depending on how the input stop (or rather the recovered articulatory trajectory) was segmented: whether it was perceived as a single segment (C), two segments (CC), or it was not heard (Ø).

Table 6.41 Confusion matrix: Segmentation of C_2 and C_3 (ratio)

		C	*CC*	*Ø*
C_2	p	**0.89**	0	0.11
	p'	**0.88**	0.06	0.06
	t	**0.81**	0.02	0.17
	t'	**0.81**	0.07	0.13
	mean	*0.81*	*0.04*	*0.12*
C_3	p	**0.92**	0	0.08
	p'	**0.97**	0	0.03
	t	**0.91**	0	0.09
	t'	**0.97**	0.01	0.02
	mean	*0.94*	*0.00*	*0.06*

As we can see, most of the time an average coda consonant C_2 is identified as a single segment (.81). It is missed in .12 of the cases, or it is heard as two separate segments (.04). Coronals are missed more often than labials, and the palatalized segments are more often perceived as a cluster. In these cases they are decomposed into j + C sequences.

An average onset consonant C_3 is correctly segmented .94 of the time. It is missed in .06 of the cases and is almost never perceived as two consonants. The only CC-case is /t'/ perceived as [t'j].[65] In the onset environment plain consonants are missed slightly more often than their palatalized counterparts. In general, segmentation is better in the onset position than in the coda context by .13. As we will see later, the errors in segmentation of C_2 and C_3 are related to the identification errors of the neighbouring vowels and consonants.

What types of errors do we find when the consonants are correctly segmented (as one consonant)? Table 6.42 restates the confusion patterns for C_2 and C_3 in terms of secondary articulation, place, manner, and voice. The missed segments and CC-responses are not considered here. The responses for place of articulation include labials, coronals, and velars. The label 'other' for manner stands for affricates, fricatives, and the glide /j/. The voiced responses include both voiced obstruents and /j/.

[65] This sequence is phonotactically possible in Russian (see section 2.2.2.1.3).

Table 6.42 Confusion matrix: C_2 and C_3 as features (ratio)

		paln		place			manner		voice	
		plain	pal	lab	cor	dor	stop	other	vls	vd
C_2	p	**0.85**	0.15	**0.83**	0.16	0.01	**0.97**	0.03	**0.96**	0.04
	p'	**0.66**	0.34	**0.74**	0.23	0.03	**0.87**	0.13	**0.87**	0.13
	t	**0.69**	0.31	0.16	**0.82**	0.02	**0.92**	0.08	**0.91**	0.09
	t'	0.32	**0.68**	0.07	**0.92**	0.01	**0.79**	0.21	**0.79**	0.21
	mean	*0.63*	*0.37*	*0.45*	*0.53*	*0.02*	*0.86*	*0.11*	*0.88*	*0.12*
C_3	p	**0.98**	0.02	**0.98**	0.02	0.00	**1.00**	0.00	**1.00**	0.00
	p'	0.12	**0.88**	**0.92**	0.08	0.00	**0.99**	0.01	**0.99**	0.01
	t	**0.89**	0.11	0.07	**0.93**	0.00	**0.96**	0.04	**0.99**	0.01
	t'	0.03	**0.97**	0.01	**0.99**	0.00	**0.95**	0.05	**0.96**	0.04
	mean	*0.51*	*0.49*	*0.49*	*0.51*	*0.00*	*0.95*	*0.03*	*0.99*	*0.01*

In the case of perfect recognition we would expect the following mean values for features: plain and palatalized (both .50); labial (.50), coronal (both .50) and velar (.00); stop (1.00) and other (0.00); voiceless (1.00) and voiced (.00).

As we can see from the table, in general a coda consonant, C_2, can be characterized as a plain (.63) rather than palatalized (.37). It is slightly more often perceived as coronal (.53) than labial (.45). It is predominantly identified as voiceless (.88) and as a stop (.86). In terms of palatalization, /p/ shows the lowest error rate (.15), while /p'/ is most often perceived as its plain counterpart (.66). /t/ and /t'/ have about the same confusion levels (.31-.32). There are fewer errors in terms of place: /p'/ is the most frequently confused segment (.26), and /t'/ is the least frequently misidentified (.08). Notice that the majority of the place confusions are between labials and coronals (on average .16). These two classes, however, are rarely confused with dorsals (all of them are /k/: .02).

In terms of both manner and voice the coda /p/ shows far fewer errors (.03 and .04 respectively) than /t'/ (both .21). A large number of these errors are due the confusion of palatalized consonants with /j/. Other manner errors include the identification of stops as affricates and fricatives of the same or different place.

As mentioned above, none of the listeners confused the stops with sonorants other than /j/. This suggests that manner differences between obstruents and sonorants are very robust. The exception was the confusion of stops (mostly palatalized ones) with /j/ in coda (/p'/: .12 and /t'/: .13). In these cases the palatalization gesture/feature was perceived as a separate segment and the primary articulation gesture/feature (Lips or Tongue Tip) of the stop was likely to be attributed to the following segment. These errors are similar to segmentation errors where a palatalized consonant was perceived as a sequence of /j/ + stop (.07 for /p'/ and .06 for /t'/) in coda and as the combination stop + /j/ (only .01 for /t'/) in onset.

Surprisingly, there was only one response identifying /t'/ as palatalized post-alveolar affricate /tʃ/ in coda. Although the two segments are relatively similar, they appear to be quite distinct for Russian listeners.[66] While plain consonants are sometimes confused with corresponding fricatives, our results do not show any confusion between palatalized stops and fricatives (e.g. /f'/, /s'/, or /ʃ:/).

The onset consonants show substantially less confusability in terms of all features. An average C_3 is a plain (.51) or palatalized (.49), coronal (.53) or labial (.49), voiceless (.99) stop (.95). In terms of palatalization /p'/ and /t/ show more errors (.12 and .11) than /p/ and /t'/ (.02 and .03). This is also the case do in terms of place (.08 and .07 vs. .02 and .00). None of the consonants was identified as a dorsal. Compared to the perception of palatalization in the coda, its recognition in onset position is better by .29. The difference between the two positions in terms of place is .13. Perception of manner and voice both improve by .11.

In sum, the current results provide support for the previous findings: for all four consonants, the number of errors is much higher in the coda than in the onset. Some consonants (/p'/ and /t/) are more often confused than others (/p/ and /t'/).

At the same time, the current experiment reveals new facts about confusion patterns among the stops. There is substantially more variability in perception of the coda consonants. These consonants are frequently either missed, or segmented as two consonants, or confused with a

[66] Both of these Russian consonants are categorized as /t'/ [tɕ] by Japanese speakers.

wide range of segments. Among the feature classes palatalization shows the most variation in the coda. Also coronals are more prone to confusion in terms of other features (manner and voice). The previous experiments could not provide these important results due to their methodological limitations.

6.4.3.1.2 C_1 and C_4

In this sub-section I look at the perception patterns of C_1 and C_4 in the sequence $C_1V_1C_2C_3V_2C_4V_3$. These consonants, not investigated before, present an interesting case. Neither C_1 nor C_4 is palatalized in the signal. They are separated from other consonants (C_2 or C_3) by a vowel. C_1 is phrase-initial, followed by /a/. C_4 is intervocalic, preceded by /a/ and followed by /i/ ([ɨ]). Does the quality of C_2 and C_3, and particularly their secondary articulation, affect the perception of these segments? It is quite reasonable to expect that it is not due to non-adjacency of these segments. Given this, the variation in the perception of C_1 and C_4 should be random. However, if they are influenced, we should find the variation to be dependent on the quality of either C_2, or C_3, or both (and not on the vowel, which is constant). Also, if we observe some effect, it is interesting to determine whether it affects both consonants to the same degree.

Table 6.43 presents the confusion results for onset consonants C_1 and C_4. The values for C_1 and C_4 were constant throughout the experiment, /t/ and /p/ respectively. As before, the responses are organized by secondary articulation and place.

Table 6.43 Confusion matrix: C_1 and C_4 perceived as one segment (tokens)

		plain								palatalized									Ø
		lab		cor					dor	lab	cor					jC/Cj			
input		p	f	t	d	ts	s	z	k	p'	t'	d'	ts'	tʃ	j	jp	jt	t'j	
C_1	t			**742**	1	44	3	1		1	60		2		1				8
C_4	p	**848**	8	3															4

Tables 6.44 restates the confusion matrix in terms of segmentation errors.

Table 6.44 Confusion matrix: Segmentation of C_1 and C_4 (ratio)

		C	CC	Ø
C_1	t	**0.99**	0.00	0.01
C4	p	**1.00**	0.00	0.00

As we see, both consonants are overwhelmingly perceived as single segments (.99-1.00) and very rarely missed (only .01 for C_1). None of the consonants is perceived as a sequence of segments.

Table 6.45 combines the responses in terms of features (excluding the missed instances).

Table 6.45 Confusion matrix: C_1 and C_4 as features (ratio)

		paln		place			manner		voice	
		plain	pal	lab	cor	other	stop	other	vls	vd
C_1	t	**0.93**	0.07	0.00	**1.00**	0.00	**0.94**	0.06	**1.00**	0.00
C_4	p	**1.00**	0.00	**1.00**	0.00	0.00	**0.99**	0.01	**1.00**	0.00

Based on this, /t/ (as C_1) is perceived as a plain (.93) coronal (1.00) voiceless (1.00) stop (.94). It is important to note that while there are no errors in terms of place or voice, there is more confusion with respect to palatalization (.07). In these cases /t/ is mostly perceived as its palatalized counterpart /t'/.

The C_4 consonant shows even fewer errors. /p/ is never confused with palatalized consonants (.00). Its place, manner, or voice are rarely misidentified. Comparing the perception of /p/ as C_4 with the same segment as C_3 (onset in a cluster) (Table 6.46), we notice that there is little difference in the perception of the two: both show near 100% identification.

The identification of /t/ as C_1 is somewhat different from the same segment as C_3: it shows fewer place (by .07) errors and about the same percentage of palatalization mistakes (less by .04). There is no difference between the two in terms of manner and voice. Are palatalization errors for /t/ as C_1 random or dependent on the palatalized quality of the following consonants (C_2 or C_3)? The analysis of tokens with a palatalized C_1 shows the following. In utterances with clusters /t/ is confused with /t'/ when followed by 2 plain consonants (C_2 and C_3 as CC) only 3% of the time. At the same time, when it is followed by one or two palatalized consonants (C'C, CC', or C'C'), it is perceived as palatalized in 11% of all cases. In utterances with single segments, the following palatalized coda consonants affect the perception of /t/ as C_1 (/p'/ .14 /t'/ .08), while the following onset consonants do not show any influence (both .00). The particular environments in which C_1 is most susceptible to palatalization are discussed in section 6.1.6.

In sum, the results show that perception errors are not limited to the consonants that have been our main focus, C_2 and C_3, but are also found in the identification of the preceding segment C_1. This consonant tends to be perceived as palatalized when at least one of the following non-adjacent, consonants is palatalized. This influence is regressive in direction. The opposite, palatalization of the following consonant, C_4, does not occur in our data. This, however, can also be due to the fact that /p/ (C_4) is less susceptible to palatalization than /t/ (C_1). Future work should test perception of C_1 and C_4 alternating their quality between /p/ and /t/.

6.4.3.1.3 V_1 and V_2

Having investigated the perception of four consonants, I turn to the vowels that precede and follow the cluster C_2C_3. Both of the vowels, V_1 and V_2, are represented by /a/, which is a low central vowel in Russian. There are only two front vowels in Russian: /i/ and /e/ [ɛ]. The sequences of vowels /ia/, /ie/, /ai/ and /ei/ are also possible, however, with the latter two allowed only when /i/ is stressed (Avanesov 1972). In this study I assume that any perception of preceding or following /a/ as a front vowel or a sequence of vowels where one of them is front can be attributed to the influence of palatalized consonants. We may also expect the opposite result: backing of /a/ to /o/ [ɔ] or /i/ [ɨ] after plain (or rather, velarized) consonants. We should keep in mind, however, that vowel reduction can also affect the choice of the vowel. Recall that in a number of tokens (the sequences of the type *ta Capy*) /a/ is substantially reduced (it belongs to the carrier phrase *e*[ta] __ *opjat'*). Although a reduced /a/, [ʌ] or [ə], is orthographically represented in Russian as *a* or *o* (Avanesov 1972), we may expect some variation in responses.

Table 6.46 presents the confusion results for the vowels V_1 and V_2. The value for both of them was /a/ throughout the experiment. The responses are divided into 'back' and 'front'. The sequence [ai] is counted together with front vowels. The missed segments are labeled as Ø.

Table 6.46 Confusion matrix: V_1 and V_2 as segments (tokens)

	back				front			
input	a	o	u	y	e	i	ai	Ø
V_1 a	**789**	3			62		3	7
V_2 a	**858**	2					1	3

Table 6.47 shows the perception of vowels in terms of the classes front and back (excluding the missed vowels).

Table 6.47 Confusion matrix: V_1 and V_2 as features (ratio)

		back	front
V_1	a	**0.92**	0.08
V_2	a	**1.00**	0.00

V_1 is identified as back .92 of the time, and as front in .08 of all the cases. The error responses consist mostly of mid front /e/ [ɛ] and a few tokens of the vowel sequence /ai/. At the same time, V_2 is always perceived as back (1.00; there is only one front token).

Are the errors in perception of V_1 conditioned by the following consonants? The analysis of /ai/ tokens shows that all three occurrences of this sequence are found before /p'/. The /e/ tokens, however, are almost evenly distributed between all following consonants with a somewhat higher number before /t'/ (_p: 15, _p': 14, _t: 12 _t': 21). A more detailed analysis of the environments shows that more than half (53%) of these tokens are found in sequences with single onset consonants (*ta Capy*), i.e. where it was more reduced in duration. Apparently, the degree of stress plays a certain role in the perception of V_1 in these utterances.

In sum, the results reveal that perception errors affect not only consonants (C_1, C_2, and C_3), but possibly also the preceding vowel V_1. The tendency of this segment to front before palatalized consonants should be further confirmed in non-native perception. It is obvious, however, that we do not find a progressive spread of palatalization: V_2 is never affected.

Among all the first four segments of the utterance, the coda C_2 is the most vulnerable to misidentification. It shows a wide range of confusion patterns that are not limited to palatalization alone. Its perception is strongly influenced by the quality of the following consonant. The influence in the opposite (of C_2 on C_3) direction is much less common. The palatalized quality can 'spill over' to the preceding vowel and even to the non-adjacent preceding consonant. However, it does not affect the following segments (V_2 and C_4).

6.4.3.2 Environments

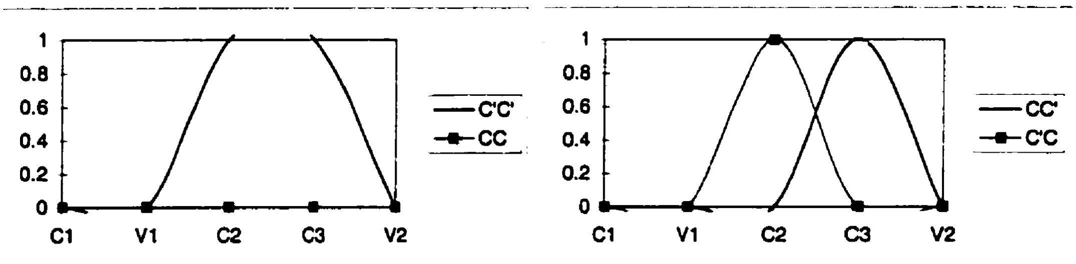
As we have seen in previous sections, the palatalized gesture of C_2 or C_3 can be 'spread' to the preceding consonant C_1 or the preceding vowel V_1. To better understand when this happens, I compare utterances in terms of their 'degree of palatalization', i.e. how often a given utterance, and each segment of it, are perceived as palatalized.

To do this, I coded the first three consonants (C_1, C_2, and C_3) and first two vowels (V_1 and V_2) for each token as either plain (.00) or palatalized (1.00) based on the confusion matrices. C_4 and V_3 were not analyzed, since they did not show any variation. For consonants, any palatalized consonant or the sequence j + C were treated as palatalized, and thus were assigned 1.00. For vowels, front vowels /e/, /i/, or sequences /ai/ were considered palatalized. For example, a response *t'ap t'apy* for the input utterance *tap' t'apy* was coded as C_1(1.00), V_1(.00), C_2(.00), C_3(1.00), V_2(.00).

Based on this, I determined the mean degree of palatalization for each timing slot in the sequence (C_1, C_2, C_3, V_1, and V_2) for all tokens of a given utterance and the overall degree of palatalization for a given utterance. The utterances with single consonants (*ta Capy* and *taC apy*) are analyzed in the same way as the utterances with clusters, since responses showed variation in assigning the word boundary and the number of consonants. That is, for instance, the input utterance *ta t'apy* could have been parsed as *ta t'apy, tat' apy, or tat' t'apy.*[67] The corresponding scores for C_2 and C_3 would be .00-1.00, 1.00-.00, and 1.00-1.00.

I represent the degree of palatalization as curves that have the highest value (1.00) when the particular segment (C_1, V_1, C_2, C_3, or V_2) is always perceived as palatalized, and as .00 when it is always identified as plain. Figure 6.29 presents the degree of palatalization curves for four cluster types (CC, C'C', C'C, CC'), assuming the perception is perfect, i.e. the perceived degree of palatalization is the same as in the input sequences. For instance, the perceived degree of palatalization is 1.00 for C_2 and C_3 in the utterance with C'C' (Figure 6.29a) and it is .00 for these two consonants for the utterance with CC. In both cases C_1, V_1, and V_2 are .0. For the CC' sequence only C_3 has the value of 1.0 (Figure 6.29b), and for the sequence C'C it is C_2 that is at 1.0.

Further we will see that perception is far from perfect both in the overall degree of palatalization and its distribution across the segments in sequence.



a. b.

Figure 6.29 Degree of palatalization assuming perfect perception: Utterances with CC, C'C, CC', and C'C' clusters

6.4.3.2.1 Utterances with single consonants

Tables 6.48 present the mean degree of perceived palatalization for a given consonant in a given utterance with single consonants and clusters. The means are calculated for each consonant and each utterance. The table also indicates the expected degree of palatalization, based on the input. The utterances are arranged by the degree of palatalization from the highest to the lowest.

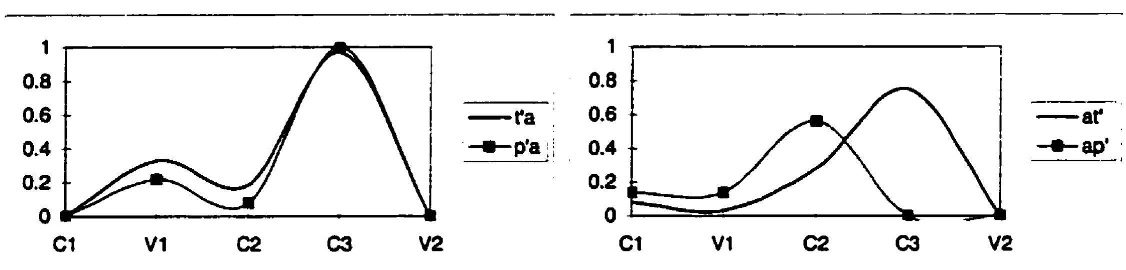
[67] This is left for further examination.

Table 6.48 Degree of perceived palatalization by segment and by utterance: Single consonants

	C_1	V_1	C_2	C_3	V_2	*Total*	*Expected*
ta t'apy	0.00	**0.33**	(0.19)	0.97	0.00	***1.50***	1.00
ta p'apy	0.00	0.22	(0.08)	**1.00**	0.00	*1.31*	1.00
tat' apy	0.08	0.03	0.28	(0.75)	0.00	*1.14*	1.00
tap' apy	**0.14**	0.14	**0.56**	(0.00)	0.00	*0.83*	1.00
ta papy	0.08	0.25	(0.08)	0.00	0.00	*0.42*	0.00
ta tapy	0.00	0.11	(0.08)	0.03	0.00	*0.22*	0.00
tap apy	0.03	0.06	0.06	(0.00)	0.00	*0.14*	0.00
tat apy	0.00	0.00	0.03	(0.00)	0.00	*0.03*	0.00
Mean	*0.04*	*0.14*	*0.17*	*0.34*	*0.00*	*0.70*	*0.50*
Expected	0.00	0.00	0.25	0.25	0.00	*0.50*	

We can see that the utterance *ta t'apy* has the highest overall score, 1.50 (Figure 6.30a). It is higher than expected based on the input. This is due to the fact that in addition to a very high palatalization of C_3, the palatalized quality is also attributed to the preceding vowel, V_1 (.33) and to C_2 (when the single consonant was perceived as a cluster).[68]

Note that there is an asymmetry in perception of coda palatalized consonants /t'/ and /p'/ (Figure 6.30b). The coda /t'/ (C_2: .28) is often interpreted as an onset consonant C_3 (.75), while this does not happen to the coda /p'/ (C_2: .56, C_3: .00). However, in the case of /p'/ in the coda we find palatalization more often spread to the preceding vowel V_1 (.14) and the preceding consonant C_1 (.14). It is interesting that the onset plain consonants tend to have a higher degree of palatalization than the coda plain consonants. The palatalized quality in these sequences is often perceived on C_2 (if misparsed), V_1, and C_1. There are substantially more confusion errors in the $C_1V_2(C_2)$ syllable, compared to the C_3V_2 syllable.



a. b.

Figure 6.30 Degree of palatalization:
Utterances with single palatalized consonants (Russian listeners)

Overall, palatalized coronals show a higher mean degree of palatalization than palatalized labials. It is the reverse for the plain consonants. Among the timing slots, C_3 has the highest palatalization rate, .34 (higher than the expected .25 by .09), while the rate for C_2 is half as high, .17 (lower by .08). It is followed by V_1 (.14) and C_1 (.04). For both of these the rate is higher than is expected based on the input. V_2 is never perceived as palatalized (.00). In general, listeners tend to hear palatalization more often than it is actually present in the input (.70 out of .50).

6.4.3.2.2 Utterances with clusters

Table 6.49 presents the mean degree of palatalization for consonants in clusters.

Table 6.49 Degree of palatalization by segment and by utterance: Consonants in clusters

	C_1	V_1	C_2	C_3	V_2	*Total*	*Expected*
tat' p'apy	**0.29**	0.11	**0.85**	0.74	0.00	*1.99*	2.00
tat t'apy	0.06	0.03	0.44	**1.00**	0.00	*1.53*	1.00
tat' t'apy	0.06	0.00	0.47	0.97	0.00	*1.50*	2.00
tap' p'apy	0.06	0.06	0.53	0.83	0.00	*1.47*	2.00
tat p'apy	0.03	0.03	0.46	0.91	0.00	*1.43*	1.00
tap p'apy	0.17	0.00	0.42	0.81	0.00	*1.39*	1.00
tap t'apy	0.14	0.00	0.06	0.92	0.00	*1.11*	1.00
tap' t'apy	0.11	0.00	0.06	0.91	0.00	*1.09*	2.00
tat' tapy	0.11	0.03	0.72	0.22	0.00	*1.08*	1.00
tap' tapy	0.25	0.00	0.36	0.25	**0.03**	*0.89*	1.00
tat' papy	0.03	0.03	0.78	0.06	0.00	*0.89*	1.00
tat papy	0.09	**0.15**	0.38	0.00	0.00	*0.62*	0.00
tap' papy	0.06	0.06	0.28	0.03	0.00	*0.42*	1.00
tap papy	0.00	0.11	0.11	0.00	0.00	*0.22*	0.00
tat tapy	0.03	0.00	0.06	0.03	0.00	*0.11*	0.00
tap tapy	0.00	0.03	0.00	0.00	0.00	*0.03*	0.00
Mean	*0.09*	*0.04*	*0.37*	*0.48*	*0.00*	0.98	*0.50*
Expected	0.00	0.00	0.50	0.50	0.00	*0.50*	

The utterance *tat' p'apy* with the hetero-organic palatalized cluster /t'p'/ has the highest overall score, 1.99. It does not, however, exceed the expected rate, 2.00. In this utterance the palatalized quality is spread over the $C_1V_1C_2C_3$ sequence with the highest rate for the coronal /t'/ as C_2 (.85). This is also the highest rate for the coda consonant C_2. The labial /p'/ of C_3 has a lower rate, .74. C_1 has the highest palatalization rate among all utterances, .29. The lowest degree

[68] It should be noted, however, that the vowel quality in the *ta Capy* utterances, as I mentioned above, may have been affected by the vowel reduction.

of palatalization is attested in the utterance *tap tapy* (.03), where the coda consonant is a plain labial and the onset consonant is a plain coronal.

The degree of palatalization is higher than expected for the utterances with CC' clusters: *tat t'apy* (by .53), *tat p'apy* (by .43), *tap p'apy* (by .39), *tap t'apy* (by .11). In these cases the secondary articulation of C_3 is also attributed to C_2. Note that the coronal /t/ is more likely to assimilate perceptually to the following consonant than /p/. In fact, the utterances with /p/ as C_2 (especially /tp/) show a higher degree of palatalization on C_1. Also we find that the assimilation of C_2 to C_1 is higher in homorganic clusters.

In addition, homorganic utterances with plain consonants, *tap papy* (by .22), *tat tapy* (by .11), as well as the sequence *tat papy* (by .62), show a higher than expected degree of palatalization. In the last sequence the preceding vowel and consonant are substantially affected (.15 and .09).

Lower than expected degree of palatalization is attested in utterances of the type C'C: *tap' papy* (by .58), *tap' tapy* (by .11), *tat' papy* (by .11). In the utterances with C_2 as /p'/ the first consonant of the utterance is more often affected. Other sequences that show lower palatalization degree than expected are those with the homorganic palatalized consonants, *tap' p'apy* (by .53), *tat' t'*apy (by .50). The lowest drop, by .91, is shown by the sequence *tap' t'apy*. /p'/ as C_2 is rarely perceived as palatalized here (.06). The palatalization is to some extent attributed to C_1 (.11). Interestingly, the results for the palatalized-palatalized labial-coronal cluster /p't'/ are the opposite of those for the plain-plain coronal-labial cluster /tp/. The latter exhibits the highest degree of palatalization attributed to the preceding vowel.

In all of the above cases the palatalization errors (either more or less palatalization than expected) primarily involve the first syllable, $C_1V_1C_2$, and mainly C_2. The syllable, C_3V_2, is rarely influenced.

Among the timing slots, C_3 has the highest palatalization rate, .48 (out of expected .50), while the rate for C_2 is lower, at .37 (lower than expected by .13). It is followed by C_1 (.09) and

V_1 (.04). V_2 is never perceived as palatalized, except in the utterance *tap' tapy* (.03). In general, listeners tend to hear palatalization substantially more often than it is actually present in the input (.98 out of .50). Overall, clusters with palatalized coronals show a higher mean degree of palatalization than palatalized labials. It is the reverse for the plain consonants.

Figure 6.31 plots all utterances with clusters (the first four segments) in terms of their degree of palatalization. They are presented by the type of the cluster: CC (ab), C'C (cd), CC' (ef), and C'C' (gh). The first column of graphs compares the patterns for labials before and after homorganic and hetero-organic plain and palatalized consonants (aceg). The second column shows coronals under the same conditions (bdfh).

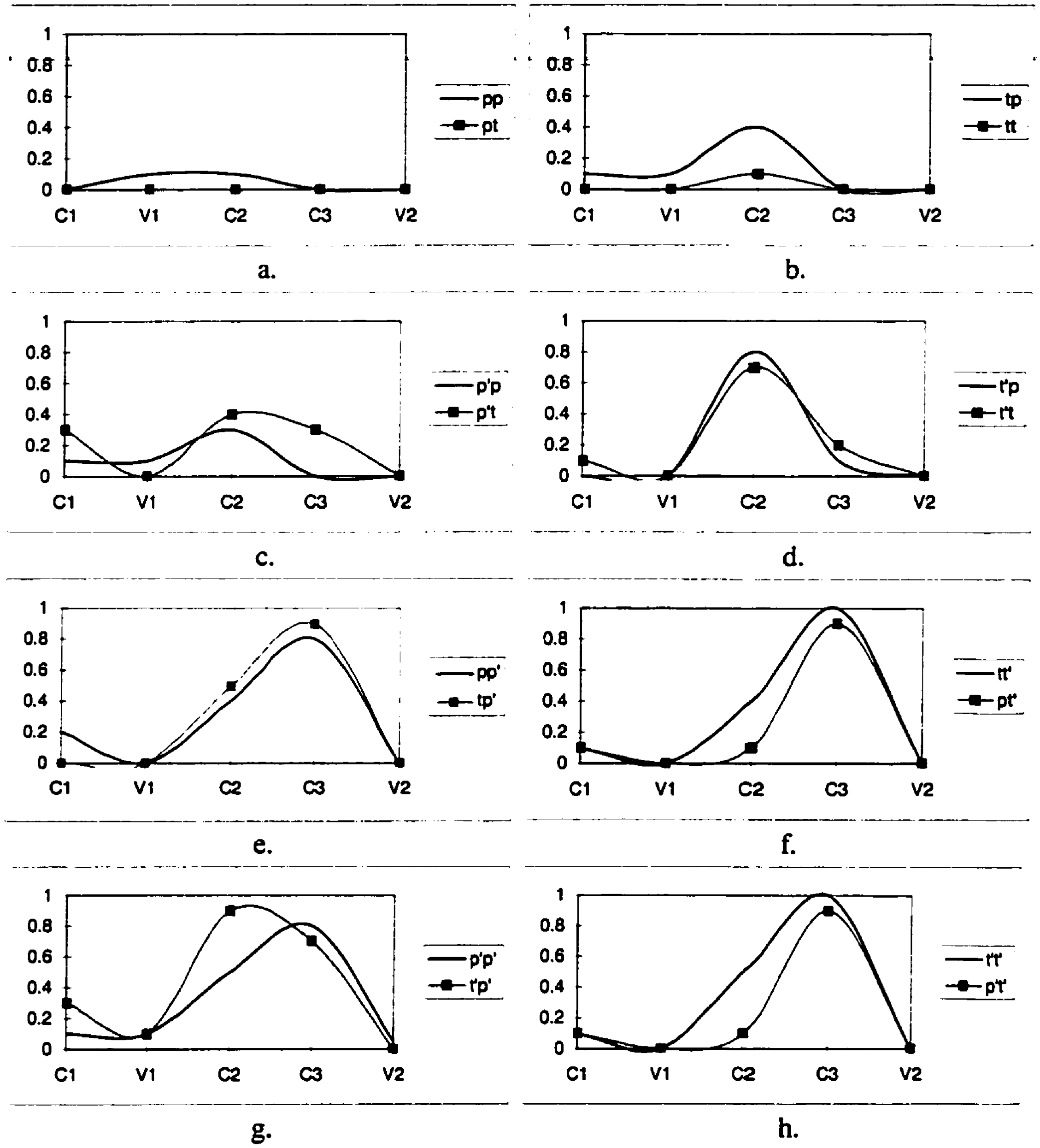


Figure 6.31 Degree of palatalization: Phrases with clusters (Russian listeners)

Comparing the two columns, we notice that labials, whether palatalized or plain, show a lower overall degree of palatalization and a stronger tendency to affect the preceding segments, compared to coronals. The same tendency is seen in Figures 6.31 efgh.

The first two rows of figures compare the sequences with CC and C'C clusters. Notice that they do differ in terms of their degree of palatalization, while the difference is less substantial in sequences with labials in coda. These figures also show that whether the following consonant is homorganic or hetero-organic is an important factor. In hetero-organic clusters C_1 shows a higher degree of palatalization (except the clusters with plain labial /p/ as C_1).

The second two rows of figures compare the utterances with CC' and C'C' clusters. The difference between the corresponding patterns is very small, particularly for clusters with C_2 as /t'/. The homorganic /pp'/ cluster is also similar to its counterpart /p'p'/. Notice that in hetero-organic clusters (/t'p'/, /pt'/, /p't'/) C_3 tends to have a lower degree of palatalization. The temporal difference between /p'p'/ and /t'p'/ is particularly obvious: the coda /t'/ attracts palatalization.

In sum, the degree of palatalization and its distribution across the utterance are rather complex and dependent on the quality of the C_2 and C_3 as well as their order. Again, we find asymmetries between coda (C_2) and onset (C_3) consonants, palatalized labials and coronals, and homorganic and hetero-organic environments. These asymmetries, however, are not only in terms of the degree of palatalization of C_2 and C_3, but also with respect to the influence exerted on the preceding segments V_1 and C_1.

6.4.4 Conclusion

The current experiment provides crucial evidence for the fact that palatalized consonants are preferred in syllable onset position and disfavored in the syllable coda context, where they are more often interpreted as plain. The palatalized quality of the coda consonant, however, is not always lost. It is attributed to the neighbouring (mostly preceding) onset consonants and vowels. Thus, the presence of palatalization not only affects the consonant in question, but also induces additional errors in the context.

In the following section we seek to confirm the current findings by testing the perception of the same utterances by Japanese listeners.

6.5 Writing task: Japanese listeners

6.5.0 Introduction

In this section I investigate perception of plain and palatalized stops by Japanese listeners. The task is the same as in the previous experiment, written response. Our main focus is similarity between Russian and Japanese subjects in recovering gestures/features.

6.5.1 Experimental setup, materials, procedure, and analysis

The design, materials, and procedure of the current experiment were the same as for Russian subjects (section 6.4.1-6.4.2). The same Japanese subjects who participated in the identification experiment discussed in section 6.1.3 were involved in this experiment. Listeners were instructed to write down these utterances in Japanese orthography (*katakana*) on answer sheets.

The analysis was also identical to the previous experiment. A total of 160 written tokens were collected (40 utterances × 2 repetitions × 20 listeners). The responses were transcribed and entered in a spreadsheet. The transcription of the Japanese tokens was checked by a native speaker. The results were not analyzed statistically due to the relatively small number of tokens and the descriptive nature of the task.

6.5.2 Results

6.5.2.1 Segments

For our purposes it was important which consonants the listeners heard when they were presented with a given stimulus, but not how they syllabified a particular consonant. The Japanese writing system does not provide one with a way of representing stops in coda other than in a homorganic cluster. Thus, the coda consonants have to be written either as /Q/ (the 'small tsu', a consonant homorganic with the following segment) or as an additional syllable. In this experiment, when the responses had the sequences *tap(papue), tapu, tappu, tappe*, and *tappu*, I assumed that the second consonant was /p/ (or /pp/).[69] Similarly, other stop sequences were treated as /p'/ (*tap(pya), tapyu, tappyu, tapi*), /t/ (*tat(ta), tatte, tattou*), and /t'/ (*tat(tya), tatyu, tatti, tattyu, tattya*).[70] The same was done with other segments (fricatives and affricates) found in the responses. In my analysis I did not distinguish between the responses where a coda consonant was represented by a single consonant or a geminate (e.g. *tap tapy* written down as *tapu tapue* or *tappu tapue*).[71]

In a few cases the responses had a word final /Q/ (the 'small tsu') that is known to represent the glottal stop in some interjections in Japanese (Vance: 1987: 12-13).

6.5.2.1.1 C_2 and C_3

Table 6.50 presents the confusion results for the coda consonant C_2 and the onset segment C_3.

Table 6.50 Confusion matrix: C_2 and C_3 perceived as one segment (tokens)

		plain											palatalized					Ø
		lab		cor						other			lab		cor			
input		p	b	f	t	d	ts	dz	s	k	?	h	p'	b'	t'	d'	h'	
C_2	p	**55**	2	2	4		1		1	7			7		3	2		
	p'	**59**	3		2					10	1		**3**	1	3	1		17
	t	**25**	2	3	10		1	2	1	12	3		3		9	2		23
	t'	14	1	1	5					9			4	2	**17**	3		**43**
total		*153*	*8*	*6*	*21*	*0*	*4*	*2*	*2*	*38*	*4*	*0*	*17*	*3*	*32*	*8*	*0*	*83*
C3	p	**81**	7	1	5							3						3
	p'	31	2		16	1		1					**39**	8	2			
	t	10	2		**73**	3		1				2			6	1		1
	t'	2	1		4								2		**69**	21	1	1
total		*124*	*12*	*1*	*98*	*4*	*0*	*2*	*0*	*0*	*0*	*5*	*41*	*8*	*77*	*22*	*1*	*5*

Overall, the four stops, /p/, /p'/, /t/, and /t'/ are perceived as 16 segments (compared to 14 for Russian subjects; section 6.4.3.1.1) out of a much wider range of possible consonants (Table

[69] Very often the Japanese subjects had difficulty dividing the utterance into two words and did not indicate word boundaries.
[70] The homorganic consonants are described as having the same secondary articulation (Akamatsu 1997).
[71] In total there were 137 geminate tokens for C_2 (39% of all the consonants representing C_2). There were also 61 geminate tokens for C4 (17% of all the consonants representing C4).

6.51).[72] Again, the choice of consonants is not random: the responses are limited to voiceless and voiced obstruents (stops, affricates, and fricatives). None of the listeners identified the stops as sonorants (nasals, liquids, or glides). It has to be noted, however, that phonotactic restrictions may be a factor here. For instance, in the Japanese sequence /a/+/i/ both vowels are usually treated as moraic (Vance 1987: 73-75), i.e. /ai/ rather than /aj/ (however, the facts of casual speech are less clear). Here I treat the response *ai* as two vowels. Note that Table 6.50 does not reflect confusion between single and geminate consonants.

Table 6.51 Japanese consonant inventory and perception of the stops /p/, /t/, /p'/, /t'/; the shaded areas represent listeners' responses

Labial		Coronal			Velar		Glottal	
plain	pal	plain	pal	pal	plain	pal	plain	pal
p	p'	t	t' [tɕ]		k	k'	[illegible]	
b	b'	d	d' [dʑ/ʑ]		g	g'		
		[ts]						
		dz [dz/z]						
[ɸ]		s	s' [ɕ]				h	h' [ç]
m	m'	n	n' [ɲ]		N [ɯ̃]			
		r [ɹ]	r' [ɹʲ]					
				j	w [ɰ]			

All the responses are divided into three groups (Table 6.52) depending on how the input stop (or rather the recovered articulatory trajectory) was segmented: whether it was perceived as a single segment (C) (or two segments), or it was not heard (∅). As we can see, the coda consonants C_2 are recovered .79 of the time and missed in .21 of the cases. Of all the consonants /p/ was always recovered, and /t'/ was missed more than any other consonant (.43). The other two segments show a segmentation error rate of about .20. The onset consonants were almost always correctly segmented (.99).

[72] Based on the analysis that treats palatalized consonants as phonemic (e.g. Akamatsu 1997).

Table 6.52 Confusion matrix: Segmentation of C_2 and C_3 (ratio)

		C	Ø
C_2	p	**1.00**	0.00
	p'	**0.83**	0.17
	t	**0.77**	0.23
	t'	**0.57**	0.43
	mean	*0.79*	*0.21*
C_3	p	**0.97**	0.03
	p'	**1.00**	0.00
	t	**0.99**	0.01
	t'	**0.99**	0.01
	mean	*0.99*	*0.01*

Table 6.53 restates the confusion patterns for C_2 and C_3 in terms of secondary articulation, place, manner, and voice. The responses for place of articulation are labeled as 'labials', 'coronals', and 'other', including dorsals and glottals. The label 'other' for manner stands for affricates and fricatives.

Table 6.53 Confusion matrix: C_2 and C_3 as features (ratio)

		paln		place			manner		voice	
		plain	pal	labial	cor	other	stop	other	vls	vd
C_2	p	**0.86**	0.14	**0.80**	0.12	0.08	**0.96**	0.04	**0.94**	0.06
	p'	**0.90**	0.10	**0.80**	0.07	0.13	**1.00**	0.00	**0.94**	0.06
	t	**0.81**	0.19	**0.44**	0.36	0.20	**0.88**	0.12	**0.93**	0.07
	t'	**0.54**	0.46	0.39	**0.46**	0.16	**0.96**	0.04	**0.89**	0.11
	mean	*0.78*	*0.22*	*0.60*	*0.25*	*0.14*	*0.95*	*0.05*	*0.93*	*0.07*
C_3	p	**1.00**	0.00	**0.92**	0.05	0.03	**0.96**	0.04	**0.93**	0.07
	p'	**0.51**	**0.49**	**0.80**	0.20	0.00	**0.99**	0.01	**0.89**	0.11
	t	**0.93**	0.07	0.12	**0.86**	0.02	**0.97**	0.03	**0.94**	0.06
	t'	0.07	**0.93**	0.05	**0.94**	0.01	**0.99**	0.01	**0.78**	0.22
	mean	*0.63*	*0.37*	*0.47*	*0.51*	*0.02*	*0.98*	*0.02*	*0.88*	*0.12*

As we can see, in general a coda consonant, C_2, can be characterized as predominantly plain (.78), labial (.60), voiceless (.93) stop (.95). In terms of palatalization /p/ and /t/ are less frequently confused with palatalized segments (.14 and .19). /p'/ is perceived as plain most of the time (.90). /t'/ is identified as plain slightly more often than palatalized (.54 vs. .46). In terms of place, there are fewer errors for labials (both .20) than for coronals (.56 for /t/ and .54 for /t'/). In these cases /t/ and /t'/ are more likely to be confused with labials (.44 and .39) than with other consonants (only the dorsal /k/: .20 and .16). The responses show much less variation in manner (the mean error rate of .05) and voice (.07). Among the four segments, /t/ is more confusable with fricatives and affricates (.12).

The onset consonants show substantially less confusability in terms of all features. An average C_3 is more plain (.63) than palatalized (.37), labial (.51) or coronal (.47) voiceless (.88) stop (.98). In terms of palatalization /p'/ shows a large number of errors (.51), while other consonants have a much lower error rate (/p/ .00, /t/ .07, /t'/ .07). Compared to the perception of palatalization in the coda, its recognition in onset is better by .28 (it is .47 for /t'/). Both /p'/ and /t/ have more place errors (.20 and .14) than the other two segments (/p/ .08 and /t'/ .06). Only .02 of responses account for other places of articulation (glottal fricatives /h/ and /h'/). None of the consonants was identified as velar. The difference between the two positions in terms of place is .24. The error rate in the perception of manner and voice do not change much (.05 vs. .02, .07 vs. .12). The onset /t'/ tends to be perceived as voiced (.22) more often than other consonants (.06-.11).

As we found in the experiment with Russian listeners, most of the errors and the widest variation in responses are found in coda. As before, the plain labial is the least affected consonant in coda. The Japanese show a higher error rate (both missing the coda stops and confusing them with other consonants) than the Russian listeners do. This can be partly explained by the difference in phonotactic restrictions as well as more subtle phonetic differences (e.g. burst duration and quality or voice onset time) between the languages.

6.5.2.1.2 C_1 and C_4

Table 6.54 presents the confusion results for consonants C_1 and C_4 and Table 6.55 groups all responses by two types of segmentation.

Table 6.54 Confusion matrix: C_1 and C_4 as segments (tokens)

	plain												palatalized						Ø
	lab			cor					other				lab		cor				
input	p	b	f	t	d	ts	dz	s	k	ʔ	g	h	p'	b'	t'	d'	s'	h'	
C_1 t	7			365	25		11	26				4	6	1	21	14	1		
C_4 p	188	50	176	3					2		1	44	5		2			3	6

Table 6.55 Confusion matrix: Segmentation of C_1 and C_4 (ratio)

		C	Ø
C_1	t	**1.00**	0.00
C_4	p	**0.99**	0.01

We can see that both consonants are correctly segmented almost all the time. Recall that the C category includes both single consonants and geminates: /p/ as C_4 was perceived as an obstruent geminate in .17 of the cases.

Table 6.56 combines the responses in terms of features (excluding the missed instances).

Table 6.56 Confusion matrix: C_1 and C_4 as features (ratio)

		paln		place			manner		voice	
		plain	pal	labial	cor	other	stop	other	vls	vd
C_1	t	**0.91**	0.09	0.03	**0.96**	0.01	**0.91**	0.09	**0.89**	0.11
C_4	p	**0.98**	0.02	**0.88**	0.01	0.11	**0.53**	**0.47**	**0.89**	0.11

Based on this, /t/ (as C_1) is perceived as a plain (.91) coronal (.96) voiceless (.89) stop (.91). The errors in terms of palatalization involve the identification of /t/ as mostly /t'/ and /d'/. /t/ is sometimes (.09-.11) confused with fricative coronals, /s/ and /z/, and voiced segments of the same place of articulation, /d/ and /d'/.

The C_4 consonant is identified as a plain (.98) labial (.88) voiceless (.89) stop (.53) or fricative (other: .47). It shows almost no palatalization errors (.02). At the same time, it is more confusable in terms of place (.12) and manner (.47): /p/ in this position is often identified as bilabial [ɸ] (*fu* and *ffu*) or glottal [h] (*he* and *hhe*), both allophones of the phoneme /h/ (Vance 1987: 21). The burst of /p/ can be transcribed as a short [ɸ], which may have been extended intervocalically. It is interesting that the Russian subjects do not seem to notice this intervocalic lenition and voicing of /p/, showing the zero error rate for both of the parameters. Note, however, that in Russian the labial fricative /f/ is labio-dental and thus more perceptually distinct from /p/.

Comparing the perception of /p/ as C_4 with the same segment as C_3 (onset in a cluster) (Table 6.15), we can notice that there is no difference in the perception of secondary articulation of the two: both show near perfect recognition. In terms of other features, and particularly, manner, the intervocalic /p/ shows more variation. The identification of /t/ as C_1 shows about the same palatalization error rate (.09) as /t/ does in the onset position in the cluster; it is better identified in terms of place (.04) and worse in terms of manner (.09) and voice (.11).

It is important to point out the difference between C_1 and C_4 with respect to errors in palatalization: while C_4 is always plain, C_1 has a tendency to be confused with its palatalized counterpart. Can this be due to the influence of the following consonants (C_2 or C_3)? In utterances with clusters, we find that /t/ is never confused with /t'/ when it is followed by 2 plain consonants (C_2 and C_3 as CC; rate .00). It is perceived as palatalized in 13% of cases when it is followed by one or two palatalized consonants (C'C: .16, CC': 11, or C'C': .13). In utterances with single segments, the following palatalized consonants do affect the perception of /t/ as C_1 (.14), while the following onset consonants show little influence (.05).

Just like Russian listeners, the Japanese subjects tend to 'spread' palatalization to the preceding non-adjacent consonant. There is no evidence for palatalizing the following consonant. There are also differences between the Russians and the Japanese in interpreting the manner of articulation and voice of prevocalic stops. This question is, however, beyond the focus of the current study.

6.5.2.1.3 V_1 and V_2

Table 6.57 presents the confusion results for the vowels, V_1 and V_2. As before, the responses are divided into back and front vowels. The sequences of vowels with /i/ are counted together with front vowels. Vowel duration, characteristic of Japanese, is ignored. None of the vowels was missed by the listeners.

Table 6.57 Confusion matrix: V_1 and V_2 as segments (tokens)

		back			front					
input		a	o	u	e	i	ai	ae	ie	ui
V_1	a	371			36		72		1	1
V_2	a	475				2		2		

Table 6.58 shows the perception of vowels in terms of the classes front and back.

Table 6.58 Confusion matrix: V_1 and V_2 as features (ratio)

		back	front
V_1	a	**0.77**	0.23
V_2	a	**0.99**	0.01

While V_2 is always identified as back (.99), V_1 is perceived as back .77 of the time, and as front in .23 of cases. In order to determine whether these are influenced by following palatalized consonants, I analyzed the error tokens. The vowel /e/ was more often found before palatalized /p'/ (12 tokens (.34)) and /t'/ (15 (.43)) than before plain segments (/p/ 3 (.09) and /t/ 5 (.14)). Most of the time the sequence of vowels /ai/ was found in the environment before /t'/ (.66). Other contexts are as follows: /p'/ .16, /t/ .11, and /p/ .06. Similarly, the environment before palatalized consonants accounted for .75 of the tokens.

It is interesting that the presence of the sequence /ai/ was found to correlate with the absence of C_2 (.75 of the /ai/ tokens). This means that the only parameter that is recovered from the input consonant /t'/ is its palatalized gesture/feature segmented as a high front vowel /i/. This vowel instead of /t'/ in the Japanese results would correspond to the /j/ in the same environment in the Russian data. This different phonemic interpretation of the palatalized gesture/feature should show in the comparison of the results for both groups.

In sum, the Japanese results provide convincing evidence that perception errors in terms of palatalization affect the preceding vowel. The listeners tend to perceive the vowel as more front when C_2 or C_3 of the stimuli are palatalized. The Japanese listeners rarely attribute palatalization to C_2, showing preference for 'spreading' it to V_1 or C_1 (and to C_3 in certain cases).

6.5.2.2 Environments

In this section I try to answer the question of how the perception of palatalization varies from utterance to utterance and to what extent this perception is similar to what we found for the Russian listeners.

6.5.2.2.1 Utterances with single consonants

Tables 6.59 present the mean degree of perceived palatalization for a given consonant in a given utterance with single consonants and clusters.

Table 6.59 Degree of perceived palatalization by utterance: Single consonants

	C_1	V_1	C_2	C_3	V_2	*Total*	*Expected*
ta t'apy	**0.30**	**0.60**	(0.15)	**0.90**	**0.05**	*2.00*	1.00
ta p'apy	0.20	0.35	(0.35)	**0.90**	**0.05**	*1.85*	1.00
tat' apy	0.05	0.55	**0.35**	(0.75)	0.00	*1.70*	1.00
tap' apy	0.00	0.35	0.00	(0.05)	0.00	*0.40*	1.00
tap apy	0.05	0.10	0.00	(0.00)	0.00	*0.15*	0.00
tat apy	0.00	0.00	0.05	(0.10)	0.00	*0.15*	0.00
ta papy	0.05	0.05	(0.05)	0.00	0.00	*0.15*	0.00
ta tapy	0.10	0.05	(0.00)	0.00	0.00	*0.15*	0.00
Mean	*0.09*	*0.26*	*0.12*	*0.34*	*0.01*	*0.82*	*0.50*
Expected	0.00	0.00	0.25	0.25	0.00	*0.50*	

We can see that the utterance *ta t'apy* has the highest overall score, 2.00 (Figure 6.32a). It is twice as high as expected based on the input. It is followed by the other palatalized onset sequence *ta p'apy* (1.85). Both of the utterances show a spread of palatalization over all four segments. The preceding vowel is particularly affected (.60 and .35). Notice that the coda consonant of *tat' apy* tends to be perceived as a palatalized onset (Figure 6.32b). To the contrary, the coda /p'/ is never, or almost never, perceived as palatalized, either in coda, or in onset, while some palatalized quality is attributed to the preceding vowel (.35). The lowest degree of palatalization is attested for *ta tapy* (.15), which has, surprisingly, a relatively high palatalization rate of C_1 (.10). In general, we find that the degree of palatalization is higher for all utterances, except for *tap' tapy*. It is particularly high for the sequences with onset palatalized consonants.

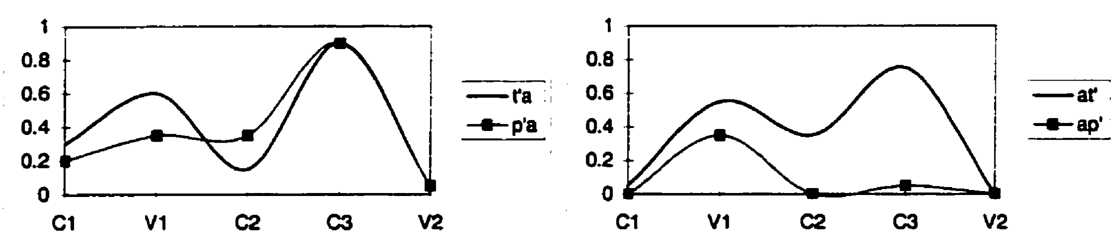


Figure 6.32 Degree of palatalization:
Utterances with single palatalized consonants (Japanese listeners)

Among the timing slots, C_3 has the highest palatalization rate, .34 (higher than expected by .09), while the rate for C_2 is almost three times lower .12 (lower than expected by .13). The rate for V_1 is higher than for C_2 (.26). It is lower for C_1 (.09) and is at its lowest for V_2 (.01). There is substantially more confusion errors in the $C_1V_2(C_2)$ syllable, compared to the C_3V_2 syllable.

In general, listeners tend to hear palatalization more often than it is actually present in the input (.82 out of .50).

6.5.2.2.2 Utterances with clusters

Table 6.22 presents the mean degree of palatalization for consonants in clusters.

Table 6.60 Degree of palatalization by utterance: Consonants in clusters

	C_1	*V_1*	*C_2*	*C_3*	*V_2*	*Total*	*Expected*
tat' t'apy	0.10	0.50	0.45	**1.00**	0.00	**2.05**	2.00
tat' p'apy	**0.25**	0.35	**0.50**	0.40	0.00	**1.50**	2.00
tat t'apy	0.05	0.15	0.45	0.85	0.00	**1.50**	1.00
tap t'apy	0.15	0.00	0.25	0.90	0.00	**1.30**	1.00
tat' papy	**0.25**	0.35	0.25	0.40	0.00	**1.25**	1.00
tap' t'apy	0.00	0.05	0.15	0.95	0.00	**1.15**	2.00
tap' p'apy	0.15	0.25	0.20	0.40	**0.05**	**1.05**	2.00
tap p'apy	0.10	0.15	0.40	0.35	0.00	**1.00**	1.00
tat p'apy	0.15	0.10	0.30	0.40	0.00	**0.95**	1.00
tat' tapy	0.15	**0.65**	0.05	0.05	0.00	**0.90**	1.00
tap' tapy	0.20	0.10	0.05	0.30	0.00	**0.65**	1.00
tat papy	0.00	0.25	0.05	0.00	0.00	**0.30**	0.00
tap' papy	0.05	0.15	0.00	0.00	0.00	**0.20**	1.00
tat tapy	0.00	0.10	0.00	0.00	0.00	**0.10**	0.00
tap papy	0.00	0.00	0.05	0.00	0.00	**0.05**	0.00
tap tapy	0.00	0.05	0.00	0.00	0.00	**0.05**	0.00
Mean	*0.10*	*0.20*	*0.20*	*0.38*	*0.00*	***0.88***	*0.50*
Expected	0.00	0.00	0.50	0.50	0.00	*0.50*	

As we would expect, the degree of palatalization in utterances is in general distributed in the following order (from highest to lowest): C'C', CC', C'C, CC. However, the place of articulation and order of the consonants is an important factor here.

The utterance with two palatalized coronals *tat' t'apy* shows the highest overall score, 2.05, which is slightly above the expected rate. It is followed by *tat' p'apy*, the utterance with the hetero-organic palatalized cluster /t'p'/ (1.50). In both utterances all four first segments have a high degree of palatalization with V_1 reaching .50 (for *tat' t'apy*) and C_1 up to .25 (*tat' p'apy*). In the latter utterance /t'/ has the highest degree for a coda consonant, C_1 (.50). Notice, however, that it is only half of the expected rate. /p'/ as C_2 in the utterance has the degree of .40, which is lower by .60 than the input. The lowest degree of palatalization is attested in the utterance *tap tapy* (.05), where the coda consonant is a plain labial and the onset consonant is a plain coronal.

The degree of palatalization is higher than expected for the utterances with CC' clusters where the second consonant is /t'/: *tat t'apy* (by .50), *tap t'apy* (by .30). The same holds for *tat' papy* (by .25) and *tat papy* (by .30). In the former the palatalized quality is almost evenly spread over the $C_1V_1C_2C_3$ sequence, with the following pain /p/ affected to the degree of .40. The utterance *tap' tapy* also presents a peculiar case: palatalized quality is perceived more on the preceding (C_1, 20) and the following (C_3, .30) coronals than on the actual palatalized segment, /p'/ (C_2, .05). The utterance *tat' tapy* shows the highest degree of palatalization on the preceding vowel, V_1 (.65), while the originally palatalized consonant /t'/ has the rate of only .05.

Lower than expected palatalization degree is attested in the utterances with C_1 as /p'/: *tap' p'apy* (by .95), *tap' t'apy* (by .85), *tat' p'apy* (by .50), *tap' papy* (by .20), and *tap' tapy* (by .35). In all of these utterances preceding segments often 'take over' some of the palatalization of /p'/.

Figure 6.33 plots all utterances with clusters (the first four segments) in terms of their degree of palatalization. As we saw before, labials, whether palatalized or plain, show a lower overall degree of palatalization. In these cases palatalization tends to be distributed more evenly among first three or four segments.

CC and C'C clusters differ in terms of their degree of palatalization. However, the difference is much greater for the sequences with labials in coda. Whether the following consonant is homorganic or hetero-organic is very important: hetero-organic clusters in general show a higher degree of palatalization which is more spread out temporally.

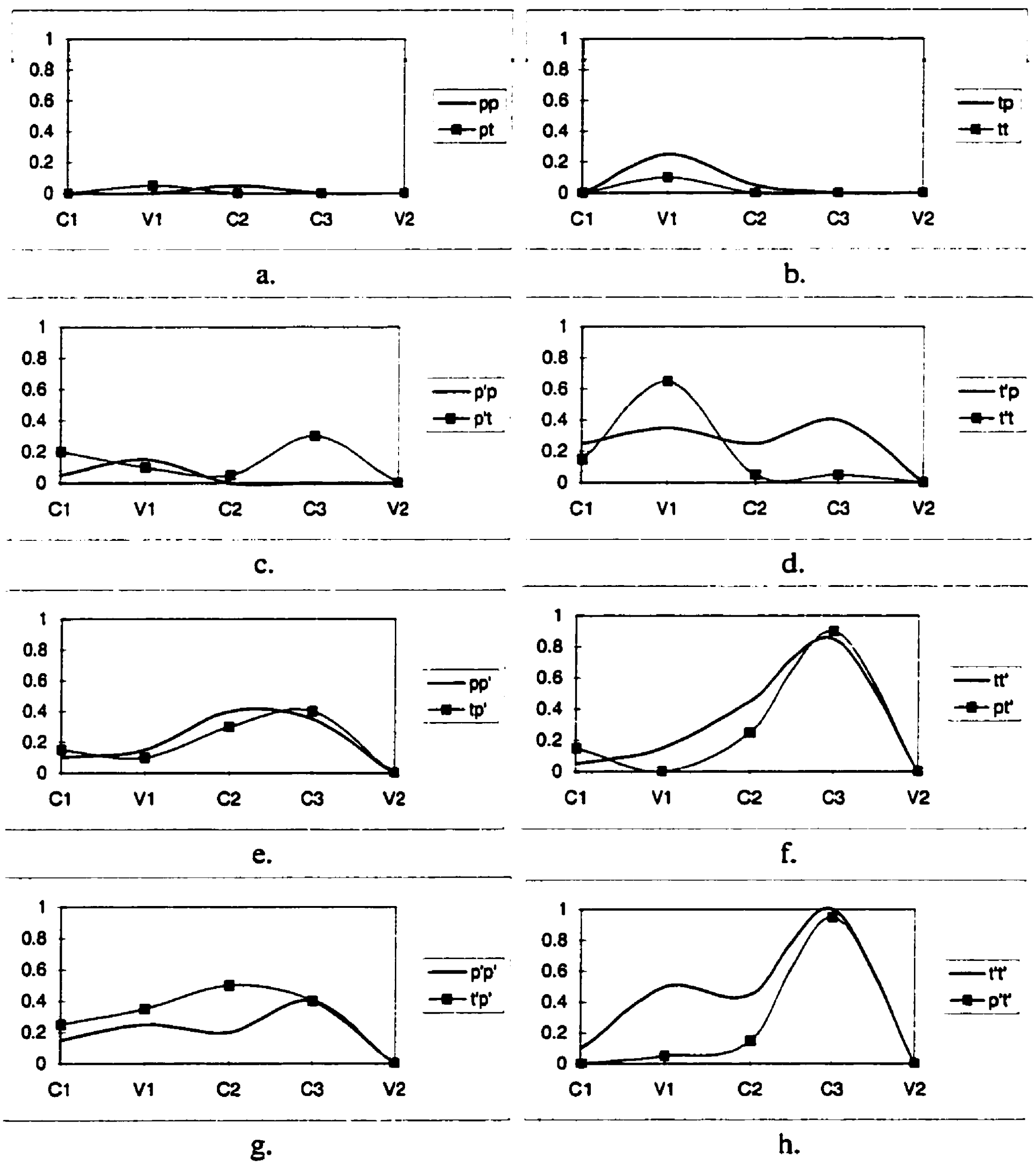


Figure 6.33 Degree of palatalization: Phrases with clusters (Japanese listeners)

The difference between CC' and C'C' clusters is less obvious, mostly showing in homorganic clusters, where V_2 rather than C_2 tends to have a higher degree of palatalization. Recall that in these sequences /at'/ was most often perceived as /ai/.

As we found in the Russian results, the perception of palatalization is not temporally restricted to the originally palatalized consonants, but rather to a sequence of several segments. These segments, or timing slots, either favour or disfavour palatalization. Onset consonants are the most likely 'sites' for the palatalized gesture/feature, while coda consonants are the least likely. Also, coronals are more likely to be palatalized than labials. The palatalized coda consonants, and particularly palatalized labials, are often treated as plain. Instead the palatalization is heard on the preceding segments.

6.5.3 Discussion

The overall degree of palatalization for consonants in clusters as perceived by both Japanese and Russian listeners is presented in Figure 6.33.

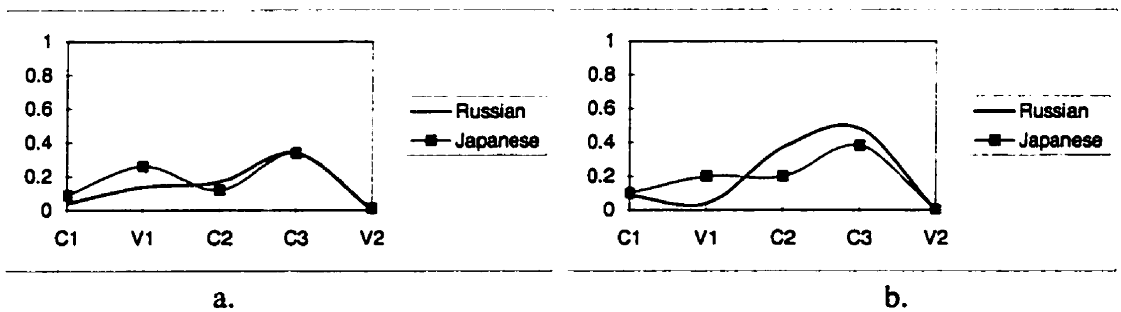


Figure 6.33 Overall degree of palatalization: Japanese and Russian listeners;
a. Utterances with single consonants; b. Utterances with consonants in clusters

Both groups of listeners have exactly the same values for the onset C_3 consonant and the following vowel V_2. However, the Japanese attribute more palatalization to the preceding vowel, V_1, than to the coda consonant, C_2. For both groups C_1 is the third in the degree of palatalization.

As we can see in Figure 6.33b the Russians and Japanese are very similar in assigning the degree of palatalization to C_1, C_3, and V_2. In fact, the values for C_1 and C_2 are the same. The value for C_3 is lower for the Japanese. However, they show a rather different treatment of C_2 and V_1. While the Russian listeners tend to hear more palatalization on C_2, the Japanese subjects show a preference for V_1. This can be partly explained by the different segmentation of the palatalized gesture/feature, which is influenced by the phonemic inventory of each language. This can also be due to the complete phonotactic restriction on coda obstruents in Japanese, while the Russian listeners are familiar with these clusters.

6.6 Conclusion

The results of the writing task from both Russian and Japanese subjects provide additional evidence supporting the earlier findings, but they also show that palatalization can be interpreted by listeners in a variety of ways. When presented with a stimulus with a palatalized consonant, a listener may correctly segment the input as a palatalized consonant or s/he may also attribute palatalization to the neighbouring (mostly preceding) consonants and vowels. A consonant in the coda environment is particularly hard to identify in terms of palatalization (but also in terms of other features). It is more often confused with the following segment. At the same time, its palatalized quality may be attributed to the preceding vowel and consonant.

The results of all perceptual experiments, summarized below (Tables 6.61-6.64), will be used in correlations between articulation and perception in the following chapter.

Table 6.61. Results of perceptual experiments: Russian subjects, phoneme identification in clean stimuli (RCI)

a. Single consonants

Stimuli	Correct identification							Perceived as					
	all	p	p'	t	t'	pal	place	p	p'	t	t'	pal	cor
ta papy	0.99	0.99				0.99	1.00	0.99	0.01	0.00	0.00	0.01	0.00
ta p'apy	0.98		0.98			1.00	0.98	0.00	0.98	0.00	0.02	1.00	0.02
ta tapy	0.98			0.98		0.98	1.00	0.00	0.00	0.98	0.02	0.02	1.00
ta t'apy	1				1.00	0.99	1.00	0.00	0.00	0.01	0.99	0.99	1.00
tap apy	0.8	0.80				0.92	0.88	0.82	0.06	0.10	0.02	0.08	0.12
tap' apy	0.38		0.38			0.57	0.69	0.33	0.37	0.10	0.20	0.57	0.31
tat apy	0.97			0.97		0.99	0.98	0.02	0.00	0.97	0.01	0.01	0.98
tat' apy	0.83				0.83	0.84	0.98	0.01	0.01	0.15	0.83	0.84	0.98

b. Onset consonants in clusters

Stimuli	Correct identification							Perceived as					
	all	p	p'	t	t'	pal	place	p	p'	t	t'	pal	cor
tap papy	0.98	0.98				0.99	0.98	0.98	0.00	0.01	0.01	0.01	0.02
tap' papy	0.95	0.95				0.94	0.99	0.94	0.05	0.00	0.01	0.06	0.01
tap p'apy	0.98		0.98			0.99	1.00	0.01	0.99	0.00	0.00	0.99	0.00
tap' p'apy	0.98		0.98			1.00	0.98	0.00	0.98	0.00	0.02	1.00	0.02
tap tapy	0.98			0.98		1.00	0.97	0.03	0.00	0.97	0.00	0.00	0.97
tap' tapy	0.82			0.82		0.85	0.97	0.03	0.00	0.82	0.15	0.15	0.97
tap t'apy	1				1.00	1.00	1.00	0.00	0.00	0.00	1.00	1.00	1.00
tap' t'apy	0.99				0.99	1.00	0.99	0.00	0.01	0.00	0.99	1.00	0.99
tat papy	0.9	0.90				0.98	0.91	0.90	0.01	0.08	0.01	0.02	0.09
tat' papy	0.92	0.92				0.96	0.96	0.93	0.03	0.03	0.01	0.04	0.04
tat p'apy	1		1.00			1.00	1.00	0.00	1.00	0.00	0.00	1.00	0.00
tat' p'apy	0.91		0.91			0.97	0.93	0.02	0.91	0.01	0.06	0.97	0.07
tat tapy	1			1.00		1.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00

c. Coda consonants in clusters

Stimuli	Correct identification							Perceived as					
	all	p	p'	t	t'	pal	place	p	p'	t	t'	pal	cor
tap papy	0.75	0.75				0.95	0.95	0.9	0.05	0.05	0	0.05	0.05
tap' papy	0.34		0.34			0.51	0.84	0.42	0.42	0.07	0.09	0.51	0.16
tap p'apy	0.35	0.35				0.43	0.87	0.35	0.52	0.08	0.05	0.57	0.13
tap' p'apy	0.72		0.72			0.79	0.91	0.19	0.72	0.02	0.07	0.79	0.09
tap tapy	0.9	0.90				0.84	0.87	0.75	0.12	0.09	0.04	0.16	0.13
tap' tapy	0.42		0.42			0.45	0.83	0.49	0.34	0.06	0.11	0.45	0.17
tap t'apy	0.68	0.68				0.73	0.87	0.68	0.19	0.05	0.08	0.27	0.13
tap' t'apy	0.37		0.37			0.55	0.79	0.41	0.38	0.04	0.17	0.55	0.21
tat papy	0.52			0.52		0.68	0.82	0.15	0.03	0.53	0.29	0.32	0.82
tat' papy	0.88				0.88	0.89	0.94	0.05	0.01	0.06	0.88	0.89	0.94
tat p'apy	0.41			0.41		0.51	0.69	0.1	0.21	0.41	0.28	0.49	0.69
tat' p'apy	0.86				0.86	0.94	0.9	0.02	0.08	0.04	0.86	0.94	0.9
tat tapy	0.72			0.72		0.88	0.82	0.16	0.02	0.72	0.1	0.12	0.82
tat' tapy	0.53				0.53	0.67	0.72	0.13	0.15	0.2	0.52	0.67	0.72
tat t'apy	0.16			0.16		0.26	0.83	0.1	0.07	0.16	0.67	0.74	0.83
tat' t'apy	0.86				0.86	0.88	0.93	0.05	0.02	0.07	0.86	0.88	0.93

Table 6.62. Results of perceptual experiments: Russian subjects, phoneme identification in noisy stimuli (RNI)

a. Single consonants

Stimuli	Correct identification							Perceived as					
	all	p	p'	t	t'	pal	place	p	p'	t	t'	pal	cor
ta papy	0.73	0.73				0.93	0.78	0.75	0.03	0.18	0.04	0.07	0.22
ta p'apy	0.25		0.25			0.85	0.34	0.08	0.26	0.07	0.59	0.85	0.66
ta tapy	0.29			0.29		0.71	0.47	0.42	0.10	0.28	0.19	0.29	0.47
ta t'apy	0.77				0.77	0.88	0.86	0.04	0.10	0.08	0.78	0.88	0.86
tap apy	0.75	0.75				0.80	0.91	0.75	0.16	0.05	0.04	0.20	0.09
tap' apy	0.23		0.23			0.42	0.72	0.49	0.23	0.09	0.19	0.42	0.28
tat apy	0.14			0.14		0.71	0.30	0.57	0.13	0.14	0.16	0.29	0.30
tat' apy	0.12				0.12	0.24	0.29	0.59	0.12	0.17	0.12	0.24	0.29

b. Onset consonants in clusters

Stimuli	Correct identification							Perceived as					
	all	p	p'	t	t'	pal	place	p	p'	t	t'	pal	cor
tap papy	0.64	0.64				0.75	0.70	0.64	0.06	0.11	0.19	0.25	0.30
tap' papy	0.52	0.52				0.83	0.62	0.56	0.06	0.27	0.11	0.17	0.38
tap p'apy	0.18		0.18			0.75	0.32	0.14	0.18	0.11	0.57	0.75	0.68
tap' p'apy	0.13		0.13			0.63	0.19	0.09	0.10	0.28	0.53	0.63	0.81
tap tapy	0.32			0.32		0.88	0.43	0.56	0.01	0.32	0.11	0.12	0.43
tap' tapy	0.31			0.31		0.48	0.67	0.17	0.16	0.31	0.36	0.52	0.67
tap t'apy	0.76				0.76	0.83	0.86	0.06	0.08	0.11	0.75	0.83	0.86
tap' t'apy	0.66				0.66	0.75	0.79	0.12	0.09	0.13	0.66	0.75	0.79
tat papy	0.52	0.52				0.73	0.61	0.51	0.10	0.22	0.16	0.27	0.39
tat' papy	0.38	0.38				0.69	0.49	0.40	0.09	0.30	0.21	0.31	0.51
tat p'apy	0.12		0.12			0.79	0.19	0.08	0.11	0.13	0.68	0.79	0.81
tat' p'apy	0.16		0.16			0.87	0.21	0.06	0.15	0.07	0.72	0.87	0.79
tat tapy	0.28			0.28		0.62	0.53	0.35	0.12	0.26	0.26	0.38	0.53
tat' tapy	0.1			0.10		0.37	0.60	0.28	0.12	0.09	0.51	0.63	0.60
tat t'apy	0.71				0.71	0.86	0.81	0.04	0.15	0.10	0.71	0.86	0.81
tat' t'apy	0.74				0.74	0.87	0.82	0.03	0.15	0.10	0.72	0.87	0.82

c. Coda consonants in clusters

Stimuli	Correct identification							Perceived as					
	all	p	p'	t	t'	pal	place	p	p'	t	t'	pal	cor
tap papy	0.52	0.52				0.67	0.80	0.52	0.28	0.15	0.05	0.33	0.20
tap' papy	0.14		0.14			0.28	0.65	0.51	0.14	0.21	0.14	0.28	0.35
tap p'apy	0.48	0.48				0.71	0.67	0.50	0.17	0.21	0.12	0.29	0.33
tap' p'apy	0.19		0.19			0.34	0.72	0.54	0.18	0.12	0.16	0.34	0.28
tap tapy	0.55	0.55				0.73	0.75	0.56	0.19	0.17	0.08	0.27	0.25
tap' tapy	0.23		0.23			0.39	0.75	0.52	0.23	0.09	0.16	0.39	0.25
tap t'apy	0.53	0.53				0.68	0.73	0.54	0.19	0.14	0.13	0.32	0.27
tap' t'apy	0.21		0.21			0.31	0.81	0.59	0.21	0.10	0.09	0.31	0.19
tat papy	0.21			0.21		0.58	0.43	0.37	0.20	0.21	0.22	0.42	0.43
tat' papy	0.19				0.19	0.37	0.40	0.42	0.18	0.21	0.19	0.37	0.40
tat p'apy	0.13			0.13		0.55	0.36	0.42	0.22	0.13	0.23	0.45	0.36
tat' p'apy	0.27				0.27	0.44	0.41	0.41	0.17	0.14	0.27	0.44	0.41
tat tapy	0.11			0.11		0.70	0.24	0.59	0.17	0.11	0.13	0.30	0.24
tat' tapy	0.25				0.25	0.42	0.41	0.41	0.17	0.16	0.25	0.42	0.41
tat t'apy	0.18			0.18		0.59	0.41	0.41	0.18	0.18	0.23	0.41	0.41
tat' t'apy	0.18				0.18	0.30	0.27	0.61	0.12	0.09	0.18	0.30	0.27

Table 6.63. Results of perceptual experiments: Japanese subjects, phoneme identification in clean stimuli (JCI)

a. Single consonants

Stimuli	Correct identification							Perceived as					
	all	p	p'	t	t'	pal	place	p	p'	t	t'	pal	cor
ta papy	1	1.00				1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00
ta p'apy	0.92		0.92			0.94	0.96	0.04	0.92	0.02	0.02	0.94	0.04
ta tapy	0.92			0.92		0.96	0.94	0.04	0.02	0.92	0.02	0.04	0.94
ta t'apy	0.98				0.98	1.00	0.98	0.00	0.02	0.00	0.98	1.00	0.98
tap apy	0.77	0.77				0.88	0.80	0.76	0.04	0.12	0.12	0.16	0.24
tap' apy	0.2		0.20			0.40	0.74	0.54	0.20	0.12	0.20	0.40	0.32
tat apy	0.41			0.41		0.82	0.55	0.41	0.04	0.41	0.14	0.18	0.55
tat' apy	0.7				0.70	0.74	0.83	0.11	0.07	0.15	0.67	0.74	0.83

b. Onset consonants in clusters

Stimuli	Correct identification							Perceived as					
	all	p	p'	t	t'	pal	place	p	p'	t	t'	pal	cor
tap papy	0.96	0.96				0.98	0.98	0.96	0.02	0.02	0.00	0.02	0.02
tap' papy	1	1.00				1.00	0.98	0.98	0.00	0.02	0.00	0.00	0.02
tap p'apy	0.96		0.96			0.98	1.00	0.02	0.98	0.00	0.00	0.98	0.00
tap' p'apy	0.98		0.98			0.96	0.96	0.04	0.92	0.00	0.04	0.96	0.04
tap tapy	1			1.00		1.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00
tap' tapy	0.96			0.96		0.58	0.94	0.02	0.04	0.56	0.38	0.42	0.94
tap t'apy	0.96				0.96	1.00	0.94	0.00	0.06	0.00	0.94	1.00	0.94
tap' t'apy	0.92				0.92	1.00	0.96	0.00	0.04	0.00	0.96	1.00	0.96
tat papy	1	1.00				1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00
tat' papy	0.56	0.56				1.00	0.96	0.96	0.00	0.04	0.00	0.00	0.04
tat p'apy	0.94		0.94			0.96	0.98	0.02	0.96	0.02	0.00	0.96	0.02
tat' p'apy	0.96		0.96			0.96	0.96	0.04	0.92	0.00	0.04	0.96	0.04
tat tapy	0.96			0.96		1.00	0.96	0.04	0.00	0.96	0.00	0.00	0.96
tat' tapy	0.98			0.98		1.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00
tat t'apy	0.98				0.98	1.00	0.96	0.00	0.04	0.00	0.96	1.00	0.96
tat' t'apy	0.92				0.92	0.98	0.98	0.02	0.00	0.00	0.98	0.98	0.98

c. Coda consonants in clusters

Stimuli	Correct identification							Perceived as					
	all	p	p'	t	t'	pal	place	p	p'	t	t'	pal	cor
tap papy	0.7	0.70				0.76	0.86	0.72	0.14	0.04	0.10	0.24	0.14
tap' papy	0.16		0.16			0.22	0.84	0.67	0.16	0.10	0.06	0.22	0.16
tap p'apy	0.46	0.46				0.50	0.86	0.46	0.40	0.04	0.18	0.58	0.22
tap' p'apy	0.51		0.51			0.68	0.74	0.24	0.50	0.10	0.18	0.68	0.28
tap tapy	0.8	0.80				0.94	0.82	0.80	0.02	0.14	0.04	0.06	0.18
tap' tapy	0.16		0.16			0.32	0.72	0.56	0.16	0.12	0.16	0.32	0.28
tap t'apy	0.61	0.61				0.67	0.75	0.60	0.15	0.06	0.19	0.33	0.25
tap' t'apy	0.18		0.18			0.49	0.59	0.41	0.18	0.10	0.31	0.49	0.41
tat papy	0.38			0.38		0.80	0.50	0.42	0.10	0.38	0.12	0.22	0.50
tat' papy	0.4				0.40	0.49	0.47	0.40	0.13	0.11	0.36	0.49	0.47
tat p'apy	0.14			0.14		0.65	0.22	0.51	0.27	0.14	0.08	0.35	0.22
tat' p'apy	0.26				0.26	0.64	0.40	0.22	0.38	0.14	0.26	0.64	0.40
tat tapy	0.3			0.30		0.78	0.46	0.48	0.08	0.30	0.16	0.24	0.46
tat' tapy	0.36				0.36	0.46	0.58	0.32	0.10	0.22	0.36	0.46	0.58
tat t'apy	0.2			0.20		0.43	0.72	0.22	0.07	0.22	0.50	0.57	0.72
tat' t'apy	0.7				0.70	0.71	0.80	0.18	0.02	0.10	0.69	0.71	0.80

Table 6.64. Results of perceptual experiments: Russian and Japanese subjects, writing task in clean stimuli (RCW and JCW)

a. Single consonants

Stimuli	Russian subjects					Japanese subjects				
	Perceived as palatalized					Perceived as palatalized				
	C_1	V_1	C_2	C_3	Total	C_1	V_1	C_2	C_3	Total
ta papy	0.08	0.25	-0.08	0	0.42	0.05	0.05	-0.05	0	0.15
ta p'apy	0	0.22	-0.08	1	1.31	0.2	0.35	-0.35	0.9	1.85
ta tapy	0	0.11	-0.08	0.03	0.22	0.1	0.05	0	0	0.15
ta t'apy	0	0.33	-0.19	0.97	1.5	0.3	0.6	-0.15	0.9	2
tap apy	0.03	0.06	0.06	0	0.14	0.05	0.1	0	0	0.15
tap' apy	0.14	0.14	0.56	0	0.83	0	0.35	0	-0.05	0.4
tat apy	0	0	0.03	0	0.03	0	0	0.05	-0.1	0.15
tat' apy	0.08	0.03	0.28	-0.75	1.14	0.05	0.55	0.35	-0.75	1.7

b. Onset consonants in clusters

Stimuli	Russian subjects					Japanese subjects				
	Perceived as palatalized					Perceived as palatalized				
	C_1	V_1	C_2	C_3	Total	C_1	V_1	C_2	C_3	Total
tap papy	0	0.11	0.11	0	0.22	0	0	0.05	0	0.05
tap' papy	0.06	0.06	0.28	0.03	0.42	0.05	0.15	0	0	0.2
tap p'apy	0.17	0	0.42	0.81	1.39	0.1	0.15	0.4	0.35	1
tap' p'apy	0.06	0.06	0.53	0.83	1.47	0.15	0.25	0.2	0.4	1.05
tap tapy	0	0.03	0	0	0.03	0	0.05	0	0	0.05
tap' tapy	0.25	0	0.36	0.25	0.89	0.2	0.1	0.05	0.3	0.65
tap t'apy	0.14	0	0.06	0.92	1.11	0.15	0	0.25	0.9	1.3
tap' t'apy	0.11	0	0.06	0.91	1.09	0	0.05	0.15	0.95	1.15
tat papy	0.09	0.15	0.38	0	0.62	0	0.25	0.05	0	0.3
tat' papy	0.03	0.03	0.78	0.06	0.89	0.25	0.35	0.25	0.4	1.25
tat p'apy	0.03	0.03	0.46	0.91	1.43	0.15	0.1	0.3	0.4	0.95
tat' p'apy	0.29	0.11	0.85	0.74	1.99	0.25	0.35	0.5	0.4	1.5
tat tapy	0.03	0	0.06	0.03	0.11	0	0.1	0	0	0.1
tat' tapy	0.11	0.03	0.72	0.22	1.08	0.15	0.65	0.05	0.05	0.9
tat t'apy	0.06	0.03	0.44	1	1.53	0.05	0.15	0.45	0.85	1.5
tat' t'apy	0.06	0	0.47	0.97	1.5	0.1	0.5	0.45	1	2.05

c. Coda consonants in clusters

Stimuli	Russian subjects					Japanese subjects				
	Perceived as palatalized					Perceived as palatalized				
	C_1	V_1	C_2	C_3	Total	C_1	V_1	C_2	C_3	Total
tap papy	0	0.11	0.11	0	0.22	0	0	0.05	0	0.05
tap' papy	0.06	0.06	0.28	0.03	0.42	0.05	0.15	0	0	0.2
tap p'apy	0.17	0	0.42	0.81	1.39	0.1	0.15	0.4	0.35	1
tap' p'apy	0.06	0.06	0.53	0.83	1.47	0.15	0.25	0.2	0.4	1.05
tap tapy	0	0.03	0	0	0.03	0	0.05	0	0	0.05
tap' tapy	0.25	0	0.36	0.25	0.89	0.2	0.1	0.05	0.3	0.65
tap t'apy	0.14	0	0.06	0.92	1.11	0.15	0	0.25	0.9	1.3
tap' t'apy	0.11	0	0.06	0.91	1.09	0	0.05	0.15	0.95	1.15
tat papy	0.09	0.15	0.38	0	0.62	0	0.25	0.05	0	0.3
tat' papy	0.03	0.03	0.78	0.06	0.89	0.25	0.35	0.25	0.4	1.25
tat p'apy	0.03	0.03	0.46	0.91	1.43	0.15	0.1	0.3	0.4	0.95
tat' p'apy	0.29	0.11	0.85	0.74	1.99	0.25	0.35	0.5	0.4	1.5
tat tapy	0.03	0	0.06	0.03	0.11	0	0.1	0	0	0.1
tat' tapy	0.11	0.03	0.72	0.22	1.08	0.15	0.65	0.05	0.05	0.9
tat t'apy	0.06	0.03	0.44	1	1.53	0.05	0.15	0.45	0.85	1.5
tat' t'apy	0.06	0	0.47	0.97	1.5	0.1	0.5	0.45	1	2.05

Chapter 7. Articulation and perception

7.0 Introduction

I have now looked at both articulation and perception. The hypothesis set out in Chapter 1 states that the act of perception involves the recovery of articulatory gestural/featural information. The goal of this chapter is to bring together the results of Chapters 5 and 6 in order to determine the correlation between articulation, acoustics, and perception, i.e. how closely articulatory differences and their acoustic consequences are reflected in listeners' responses. Among these, the asymmetries in terms of palatalization, place, and environment are of major interest to us. This will lead us to propose general perceptual scales that reflect recoverability of palatalization in various environments.

The chapter is organized as follows. First, I examine the relation between articulation (Tongue Body and primary articulators) and perception of palatalization. Then, I consider articulatory factors (primary gestures and Tongue Body) and perception of place. Further, I discuss the relevance of other articulatory and acoustic factors (overall Tongue Body difference, release rate and burst quality) for the correct identification of consonants.

7.1 Analysis

To determine the relation between articulatory input and perception, I performed a series of correlational analyses. The various articulatory results for Subject 3 (see Tables 5.40-5.41, Chapter 5) were correlated with the results of perception under various conditions (see Tables 6.61-6.64, Chapter 6). These correlations were done using Pearson's correlation coefficient for ranked data (r_s) (cf. Surprenant & Goldstein 1998).

For convenience I use the following abbreviations to refer to various conditions and tasks performed in the perceptual experiment: RCI (Russian subjects, clean stimuli, identification task), JCI (Japanese subjects, clean stimuli, identification task), RNI (Russian subjects, noisy stimuli, identification), RCW (Russian subjects, clean stimuli, writing task), JCW (Japanese subjects, clean stimuli, writing task).

7.2 Articulation and perception of palatalization

As discussed in Chapter 5, the results of the articulatory experiment show that the production of a palatalized consonant involves the gesture of Tongue Body raising and/or fronting. Each of the palatalized segments in the input is associated with this gesture, while showing different degrees of magnitude and timing. Plain consonants are characterized by lack of Tongue Body raising and, for labials, by backing of this articulator.

7.2.1 Presence or absence of Tongue Body raising/fronting

I begin by looking at the relation between the categorical presence or absence of the palatalized gesture in the input and the perception of the corresponding segment as palatalized or plain. The palatalized gesture is assumed to be present (or intended to be present) in any pronounced consonant /p'/ and /t'/ and absent (or intended to be absent) in any segment /p/ and /t/. In general, the results are as follows. The articulatory results are correlated with perceptual responses under all five conditions. There is a difference between the identification (see section 6.1) and writing (section 6.4) tasks. The focus of the former task is presence/absence of Tongue Body gesture and perception of the corresponding segment as palatalized or plain, while the latter takes into account presence/absence of Tongue Body in the utterance in general and determines whether this is related to the reported degree of palatalization of the utterance.

Note that not all of the responses (as palatalized/plain) are correct. Rather, they indicate what articulatory information is *perceived* by listeners to be a property of the corresponding segment.

Here and in subsequent sections I first report the statistical results for each correlation (without relating the results to the overall analysis). These are followed in section 7.2.1.2 (and corresponding sections) by the linguistically relevant discussion.

7.2.1.1 Results

Table 7.1 presents the results of the correlations. The significant results are given in bold. A significant correlation of [rs = .87, p =.000] for RCI means that the presence of Tongue Body raising/fronting gesture, associated with a segment, is related to the perception of this segment as palatalized. The absence of this gesture is perceived by listeners as a correlate of a plain consonant (the correlation values for the response 'plain' are not listed, since the values for them are the opposite of the ones for 'palatalized', i.e. the correlation between the presence of Tongue Body and response 'plain' in RCI is rs = -.87, p = .000).

Table 7.1 Correlation between the Tongue Body gesture and its perception

Condition	Perceived	TB fronting/raising gesture	
		rs (40)	*p*
RCI	palatalized	**0.87**	**0.000**
JCI	palatalized	**0.79**	**0.000**
RNI	palatalized	**0.57**	**0.000**
RCW	degree of palatalizn	**0.63**	**0.000**
JCW	degree of palatalizn	**0.63**	**0.000**

This correlation was significant for palatalized consonants under all conditions (RCI [rs = .87, p = .000], JCI [rs = .79, p = .000], RNI [rs = .57, p = .000], RCW [rs = .63, p = .000], and JCW [rs = .63, p = .000].

In addition, the writing task further allowed me to determine the relation between the number (1 or 2) of intended palatalized gestures/features in the input and the number of perceived palatalized consonants. This correlation, not listed in Table 7.1, was significant for RCW [rs = .49, p = .009], but not for JCW [rs = .33, p = .082].

7.2.1.2 Discussion

The statistical correlation results show that both Russian and Japanese listeners perceive the Tongue Body raising/fronting gesture/feature as secondary palatal articulation under several conditions. The correlation is particularly high in the identification of clean stimuli, and especially in native perception. It is lower under the more complex written task, since the choice is not restricted to four segments and listeners are less familiar with the stimuli, and with noise, where some acoustic information about the gesture/feature is masked.

The results also show the relative sensitivity of Russian listeners (RCW) to the duration of Tongue Body movement: the longer duration was associated with two palatalized segments under the written task. The correlation was not significant for the Japanese listeners, who often treated a longer Tongue Body movement (two gestures) as one gesture/feature and associated it with a single segment. This is explained by the language-particular duration patterns of the Tongue Body gesture. Recall from Chapter 6 that the results of the writing task also point to the difficulty listeners of both groups had with determining the correct timing of the Tongue Body gesture/feature; it could be ascribed to one of four segments: C_1, V_1, C_2, and C_3 (in $C_1V_1C_2C_3V_2C_4V_3$). This error was common for both Russian and Japanese subjects. The choice of the segment associated with Tongue Body was also influenced by the language-particular set of consonants and vowels and the phonotactics of the language. These results suggest that the recovery of Tongue Body and its association with the primary stop gesture/feature (segmentation) are independent processes. While the first process is language-independent, the second one is largely informed by the phonological knowledge of listeners.

As we will see in section 7.3.3, vertical Tongue Body movement is also used by listeners in recovering the place of articulation with noise, a factor that can lead to confusion of plain and palatalized consonants.

7.2.2 TBy

In this section I examine how listeners use the relative Tongue Body magnitude (vertical dimension, TBy) at four points in time (VC1, VC2, CV1, and CV2) and TBy slope (during the VC and CV transitions) to determine the plain/palatalized nature of a consonant.

7.2.2.1 Results

Table 7.2 presents the results of the correlation between Tongue Body raising and its perception. A significant correlation of rs = .76, p =.000 for RCI at VC1 means that higher values of TBy at this point were consistently associated with the secondary palatal articulation of the segment. Respectively, the lower values were interpreted as denoting a plain consonant.

Table 7.2 Correlation between the Tongue Body (y) raising and its perception at four points in time

Condi-tion	Perceived as	TBy VC1		TBy VC2		TBy CV1		TBy CV2	
		rs	*p*	*rs*	*p*	*rs*	*p*	*rs*	*p*
RCI	palatalized	**0.76**	**0.000**	**0.88**	**0.000**	**0.98**	**0.000**	**0.95**	**0.000**
JCI	palatalized	**0.59**	**0.002**	**0.75**	**0.000**	**0.98**	**0.000**	**0.96**	**0.000**
RNI	palatalized	0.12	0.570	0.30	0.149	**0.85**	**0.000**	**0.84**	**0.000**
RCW	degree of palatalization	**0.68**	**0.000**	**0.80**	**0.000**	**0.81**	**0.000**	**0.76**	**0.000**
JCW	degree of palatalization	**0.57**	**0.003**	**0.76**	**0.000**	**0.81**	**0.000**	**0.79**	**0.000**

The correlation between perceived palatalization and Tongue Body magnitude was found at all four points in time for all conditions, except for RNI. The former showed significant correlation at the last two points only. For all the conditions the highest correlation results are found during the CV transition (both CV1 and CV2).

Table 7.3 shows the results of the correlation of TBy slope and perceived palatalization and place. A significant correlation of rs = .90, p =.000 for RCI (response 'palatalized') during the VC transition means that a raising movement (positive slope) of Tongue Body was associated by listeners with the secondary palatal articulation of the segment. The palatalization was correlated with the negative slope (lowering movement of Tongue Body) during the CV transition. Respectively, the negative slope during the VC transition, its positive value at the CV transition, or lack of slope whatsoever, was associated with plain segments.

Table 7.3 Correlation between the Tongue Body (y) slope during VC and CV transition and its perception

Condition	Perceived as	TBy VC slope		TBy CV slope	
		rs	*p*	*rs*	*p*
RCI	palatalized	**0.90**	**0.000**	**-0.99**	**0.000**
JCI	palatalized	**0.83**	**0.000**	**-0.98**	**0.000**
RNI	palatalized	**0.44**	**0.031**	**-0.83**	**0.000**
RCW	palatalized	**0.83**	**0.000**	**-0.83**	**0.000**
JCW	palatalized	**0.86**	**0.000**	**-0.81**	**0.000**

There was a significant correlation between perceived palatalization and VC and CV slope of the vertical Tongue Body movement at all four points under all the conditions, including RNI.

7.2.2.2 Discussion

The results of the correlations suggest that listeners attend to the relative Tongue Body magnitude (TBy) throughout the examined time interval, even at Point VC1, where the difference between Tongue Body values was found to be minimal and often not significant. The higher the value of TBy, the more likely it is that the listeners will perceive it as a palatalized gesture/feature. The lower the Tongue Body, the more likely they treat it as a property of a plain consonant. Note that this correlation is almost absolute at the CV transition (rs = .95-.99 at CV1 and CV2 for RCI and JCI), showing a greater reliance of listeners on the CV transition in identification of the plain-palatalized contrast.[73] This follows the overall distribution of articulatory differences in Tongue Body in *initial* consonants: the maximum contrast is at Point CV1. Recall, however, that the articulatory facts are more complicated. First, the difference between /t/ and /t'/ is spread out through both transitions. Second, final single labials, as pronounced by Subject 3, show more contrast at VC2 than at CV1. Third, the relative articulatory differences at Point CV2 are substantially less than at Point CV1. These facts,

[73] The higher reliance on the information at CV1 can be partly attributed to the acoustic burst right before this point. It is difficult to determine the relative importance of the burst and transition. However, the results with noise, where burst noise is masked, show almost identical results.

however, do not seem to be reflected in the correlation results, suggesting that listeners are attuned to some articulatory differences (particularly those at the CV transition), while being less sensitive to the others (the VC transition). Adding noise makes the information from the VC transition even less recoverable. Thus, the articulatory asymmetry between the closing and opening gesture movements is increased by a perceptual factor, the relative salience of VC and CV transitions (Redford & Diehl 1999 and references therein).

In addition to the relative magnitude of TBy, listeners also pay attention to the overall direction (slope) of the movement of this articulator. Interestingly, there is a relation between the VC (and CV) slope and the perception of palatalization with noise. Recall that the relative TBy magnitude during the VC transition did not play any role under this condition. This can be interpreted as suggesting that the listeners notice the overall direction of Tongue Body movement (based on the abrupt changes of formants), but they fail to determine the relative height of the articulator.

Another important point is that listeners use the TBy magnitude and slope information for determining the place of articulation of the stop with noise (CV1 and CV2 only). This is discussed in section 7.3.3.

The correlation results discussed above do not take into consideration the differences between environments. Below we will see how the perception of TBy magnitude varies in these contexts.

Since Point CV1 shows the highest correlation, I use the Tongue Body values and responses at this point to represent the single consonants. The articulatory and perceptual measurements for CV1 will be also used for onset consonants in clusters, while VC1 measurements (in the absence of CV transition) will be used for the coda consonants in clusters. To be able to plot both Tongue Body values and responses, I convert the actual TBy values to percentages of the overall Tongue Body movement range in the production of single consonants (see section 5.3.2.6). The highest value is 1.00 (i.e. 100% of the range) and the lowest value is 0.00 (0% of the range). The palatalized consonants tend to have higher values of Tongue Body (above .50) and the plain segments are characterized by lower values (below .50). Note that I use the range of TBy movement for single consonants, and the segments in clusters may have values beyond this range (i.e. 1.05, i.e. 5% higher than for a single consonant).

Figure 7.1 plots the relative TBy magnitude and the corresponding perception in single consonants. The value 1.00 for /t'a/ both in terms of Tongue Body magnitude and perception means that the initial /t'/ had the highest TBy (at Point CV1) and it was perceived as palatalized 100% of the time. Looking at the results for /pa/, we can see that the initial plain labial /p/ has the lowest TBy position (0.00) and is perceived as plain 100% of the time. The TBy values for the other two initial segments, /p'/ (in /p'a/) and /t/ (in /ta/), are either lower than 1.00 (/p'/) or higher than 0.00 (/t/).

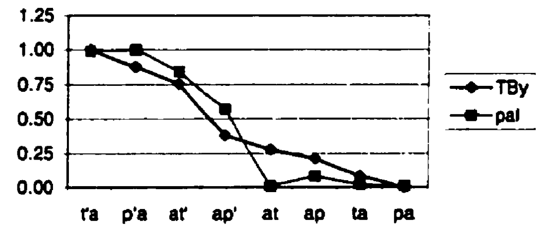
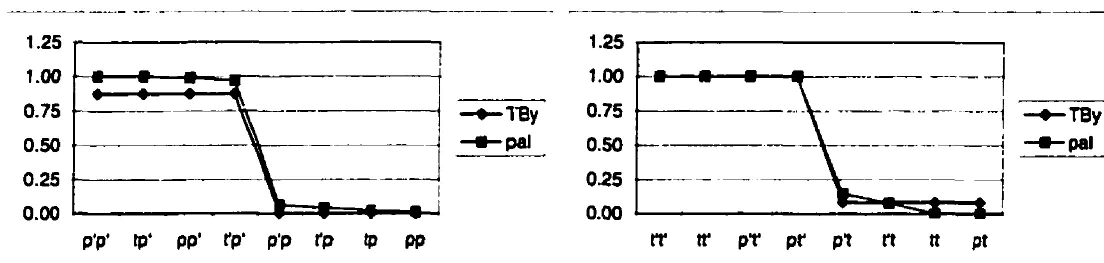


Figure 7.1 Correlation between Tongue Body (TBy) magnitude (max = 1, min = 0) and its perception (as palatalized = 1, as plain = 0), single consonants, condition RCI

Note also the TBy differences between initial and final consonants. While the overall perception of the stops generally correlates with the continuum of TBy magnitude values, there are some differences. Thus, the input palatalized and plain consonants are perceived as correspondingly 'palatalized' or 'plain' slightly more often than they should be based on the relative values of TBy. Notice that the TBy value for /p'/ is below .50, i.e. in terms of articulation it is more plain than palatalized at this point. Nevertheless, it is more often identified as palatalized. This can be explained by the listeners' use of other sources of information about the contrast: TBy values at other points in time (VC1, VC2,and CV2), as well as due to TBx values and burst. Another difference is that the perception of palatalized stops is more sensitive to the gradient change of TBy values than that of the plain segments, that show a more categorical identification.

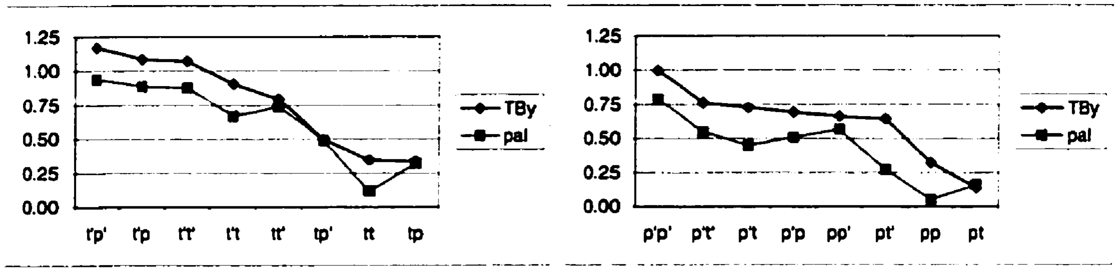
The opening movement of Tongue Body during the CV transition (together with the burst) is the only source of information about the secondary articulation of the onset consonant in clusters. Figure 7.2 shows the TBy values at CV1 (based on the single consonants) and their perception, separately for labials (a) and coronals (b) (as C_2). We can see that the responses 'plain' or 'palatalized' are highly related to the TBy values. And since the latter are relatively consistent regardless of the preceding consonant, we observe a similar consistency in responses.



a. b.

Figure 7.2 Correlation between Tongue Body (TBy) magnitude (max = 1, min = 0) and its perception (as palatalized = 1, as plain = 0), onset consonants (C_2) in clusters, labial (a) and coronal (b), condition RCI

The coda consonants in clusters do not carry information about Tongue Body release and opening movement. Instead, the VC transition contains some information about the closing movement of Tongue Body. Recall that in this environment Tongue Body values are highly sensitive to the nature of the following consonant. Figure 7.3 plots the TBy magnitude and perception responses in the preconsonantal position, separately for labials (a) and coronals (b) (as C_1). The TBY values above 1.00 for /t'p'/, /t'p/, and /t't'/ (Figure 7.3a) indicate that the articulator during the production of /t'/ in these environments is higher than during the production of single initial /t'/ (notice that the plain /t/ also shows values substantially higher than for a single /t/). Tongue Body is the highest for /t'/ followed by palatalized segments and the lowest when /t'/ is before a homorganic plain /t/. Note that the overall responses for /t'/ as 'palatalized' follow the general TBy pattern, but are somewhat lower than should be expected based on the articulatory values. The Tongue Body values for the plain /t/ show a wide range of variation, which correlates with the range of responses.



a. b.

Figure 7.3 Correlation between Tongue Body (TBy) magnitude (max = 1, min = 0) and its perception (as palatalized = 1, as plain = 0), coda consonants (C_1) in clusters, labial (a) and coronal (b), condition RCI

The correlation between TBy and the perception of plain/palatalized stops in clusters with labials as C_1 (Figure 7.3b) is similar to the coronals. The overall lower TBy values for palatalized and plain labials correspond to the higher likelihood of association of labials with plain segments.

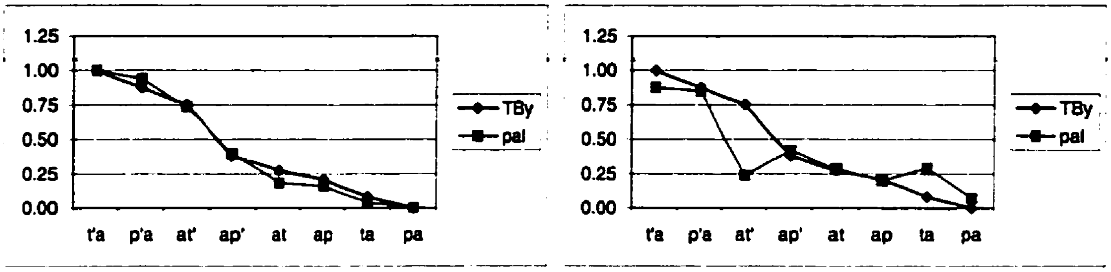
In general, the perception of palatalized and plain consonants in all environments correlates closely with the relative Tongue Body magnitude at the CV transition for single consonants and onsets in clusters, and at the VC transition for coda consonants in clusters. In the identification of the plain-palatalized contrast listeners seem to apply a fairly simple algorithm: if the recovered Tongue Body at the CV transition is 'higher' than average (i.e. about .50), the consonant is perceived as palatalized; if it is lower than average, it is categorized as 'plain'. This approach seems to be applied the same way regardless of the context. This gives very good results in the identification of initial segments or onset consonants in clusters, but leads to a number of problems in other environments.

First, the single final consonants are perceived as plain or palatalized based on the CV transition, which may not contain all the crucial information about the contrast (unless these consonants are fully re-syllabified). The VC transition (that has the information about the shifted

TBy gesture/feature of /p'/) is not consulted to the same extent. This leads to the tendency to perceive the final palatalized segments (and particularly /p'/) as plain. Second, due to the perceptual differences between CV and VC transitions, the same information is harder to recover from the latter context, where Tongue Body magnitude values are perceived as lower than they actually are (as reflected in the responses). This results in the overall higher confusability of preconsonantal segments (as shown by the perceptual experiment, discussed in Chapter 6) and a bias towards plain consonants in this environment. Third, the TBy values in preconsonantal position show greater fluctuation depending on the plain/palatalized quality of the following segment. As we saw in section 5.3.3.6.3, /p/ or /t/ before palatalized segments have a TBy magnitude around or above .50, i.e. closer to the palatalized segments. At the same time, the TBy values of the palatalized stops /p'/ and /t'/ can be as low as those of some plain segments. All this adds to the lower identification rate in this environment. Additional factors of no less importance, i.e. the quality of burst of consonants and its availability on C_1 due to degrees of lag and release rate, are considered in section 5.5.

Let us now compare the perception of single consonants under three conditions: RCI, JCI, and RNI. The correlation of TBy and perception of these consonants by native speakers (clean) is shown in Figure 7.1. Figure 7.4 plots the same articulatory input and its perception by Japanese listeners (a) and by native subjects with noise (b). The correspondence of TBy values and their labeling as 'palatalized' or 'plain' by the Japanese listeners is almost exactly one-to-one. This suggests that these subjects rely almost solely on the CV transition in their responses, ignoring the VC information altogether.[74]

[74] It also seems that they do not take advantage of the burst of final stops (e.g. /p'/).



a. b.

Figure 7.4 Correlation between Tongue Body (TBy) magnitude (max = 1, min = 0) and its perception (as palatalized = 1, as plain = 0), single consonants, conditions JCI and RNI

Perception with noise shows an increased asymmetry between initial (onset) and final (coda) positions: while initial /t'/ and /p'/ are perceived as palatalized relative to their TBy values, the same consonants in the final environment are perceived as plain. Recall that the TBy values at the VC transition do not contribute to the identification of the consonants with noise. The asymmetry may also be based on the values further on in the CV transition, at CV2, where the final /p'/ and /t'/ have much lower magnitude. It is important to point out that, while both initial /p'/ and /t'/ are perceived as palatalized, their place difference is not distinguished due to the masked burst. Both of the consonants tend to be identified as /t'/, the segment that is characterized by higher TBy. Note that the lack of perceptible burst (together with the TBy values) does not affect the perception of the palatalized quality of the initial /p'/ and /t'/, but is detrimental to these stops in the final position.

The perception of the plain/palatalized distinction in the other contexts under the conditions JCI and RNI follows the same pattern as in RCI, while showing more errors in the less salient environment, preconsonantal position.

7.2.2.3 Conclusion

In sum, there is a strong correlation between the TBy values and the perception of plain and palatalized consonants under all the conditions. This correlation is particularly significant during the CV transition. Using this measure regardless of context, however, leads to the misidentification of the plain or palatalized nature of segments in final and preconsonantal environments. This effect of the environment is even stronger in non-native perception and in native perception under adverse conditions.

7.2.3 TBx

In this section I examine how listeners use the relative Tongue Body magnitude on the horizontal dimension (TBx) at four points in time (VC1, VC2, CV1, and CV2) and TBx slope (during the VC and CV transitions) to determine the plain/palatalized nature of a consonant. Recall that this dimension of Tongue Body movement serves mainly to differentiate the contrast between plain and palatalized labials, /p/ and /p'/. A difference in TBx between the coronals /t/ and /t'/ in the production of Subject 3 is present but less substantial, being significant at CV1 (where Tongue Body for /t'/ is at a more front position). The TBx difference also contrasts the plain /p/ with all other stops (/t/, /p'/, and /t'/) that share Tongue Body fronting as a separate gesture or (/p'/) or as the movement due to Tongue Tip fronting (/t/ and /t'/). These place differences are examined separately in section 7.3.

7.2.3.1 Results

Table 7.4 presents the results of the correlation between the articulatory Tongue Body fronting and its perception. Recall that higher TBx values correspond to Tongue Body backing and lower values represent Tongue Body fronting. Thus, a negative correlation (e.g. $r_s = -.44$, $p = .032$ for RCI at VC1) means that more front (lower, in absolute terms) values of TBx at this point were consistently associated with the secondary palatalized articulation of the segment. Respectively, the more back (lower, in absolute terms) values were interpreted as denoting a plain consonant.

Table 7.4 Correlation between the Tongue Body (x) fronting and its perception at four points in time

Condition	Perceived	TBx VC1		TBx VC2		TBx CV1		TBx CV2	
		rs	*p*	*rs*	*p*	*rs*	*p*	*rs*	*p*
RCI	palatalized	**-0.44**	**0.032**	**-0.69**	**0.000**	**-0.89**	**0.000**	**-0.90**	**0.000**
JCI	palatalized	-0.34	0.110	**-0.56**	**0.005**	**-0.91**	**0.000**	**-0.92**	**0.000**
RNI	palatalized	-0.17	0.418	-0.31	0.136	**-0.85**	**0.000**	**-0.85**	**0.000**
RCW	degree of palatalizn	**-0.47**	**0.022**	**-0.63**	**0.001**	**-0.75**	**0.000**	**-0.77**	**0.000**
JCW	degree of palatalizn	-0.22	0.294	**-0.55**	**0.006**	**-0.72**	**0.000**	**-0.70**	**0.000**

The significant correlation between perceived palatalization and TBx magnitude was found at all four points in time for Russian listeners under two conditions (RCI and RCW), at the last three points for the Japanese subjects (JCI and JCW), and at the last two points with noise for Russian subjects (RNI).

The results of the correlation between the TBx slope and perception are given in Table 7.5.

Table 7.5 Correlation between the Tongue Body (x) slope and its perception at four points in time

Condition	Perceived as	TBx VC slope		TBx CV slope	
		rs	*p*	*rs*	*p*
RCI	palatalized	**0.65**	**0.001**	**-0.74**	**0.000**
JCI	palatalized	**0.54**	**0.007**	**-0.77**	**0.000**
RNI	palatalized	0.31	0.142	**-0.73**	**0.000**
RCW	palatalized	**0.57**	**0.003**	**-0.36**	**0.002**
JCW	palatalized	**0.58**	**0.003**	**-0.66**	**0.000**

There was a significant correlation between perceived palatalization and VC and CV slope of the horizontal Tongue Body movement at all four points under all conditions, except for RNI. In the former case the correlation was significant only for the CV slope.

7.2.3.2 Discussion

The results indicate that when identifying palatalization based on TBx, listeners rely on the information provided by the CV transition to an even greater degree than they do when using the TBy values. For all conditions the correlation values for CV1 and CV2 are from .70 to .92. The more front the recovered Tongue Body gesture/feature is, the more likely the corresponding consonant is perceived as palatalized. This, however, involves an additional task of differentiating between the consonants that share Tongue Body fronting: /p'/ and /t'/ vs. /t/. As discussed in section 7.2.3, this may be an additional source of confusion between plain and palatalized consonants.

The direction of the slope of TBx also correlates with the plain/palatalized and coronal/labial responses. Thus, the fronting movement of Tongue Body is interpreted as both palatalized and coronal, while the backing movement is associated with the labels 'plain' and labial'. This, again, contradicts the fact that the Tongue Body is also fronted during the production of /t/, which is both plain and coronal.

How does the perception of TBx movement vary in different environments and under various conditions? Below I examine the correlations between TBx and its identification as plain or palatalized. Since this parameter is mostly used to differentiate /p/ and /p'/, I limit the discussion to these two segments.

As in the case of TBy, I use the correlations at Point CV1 for the single consonants and onset consonants in clusters. VC1 measurements are used for the coda consonants in clusters. The actual TBx values are converted to percentages of the overall horizontal Tongue Body movement range observed in the production of single consonants (see section 7.2.2) by Subject 3. The most front value is 1.00 (i.e. 100% of the range) and the most back value is 0.00 (0% of the range). The palatalized consonant /p'/ tends to have more front values of Tongue Body (above .50), while the plain labial /p/ is characterized by lower values (below .50).

Figure 7.5 plots the relative TBx magnitude and the corresponding perception in single consonants. The value 1.00 for /p’a/ both in terms of Tongue Body magnitude and perception means that the initial /p'/ has the most front Tongue Body (at Point CV1) and it is perceived as palatalized 100% of the time. The results for /pa/ are exactly the opposite: the initial plain labial /p/ has the most back TBx position (0.00) and is perceived as plain 100% of the time. The perception of the same segments (and especially /p’/) in final position correlates with the relatively more back or more front TBx values of these segments. The higher than expected value of /p’/ can be attributed to the Tongue Body information and the use of the VC transition in its identification.

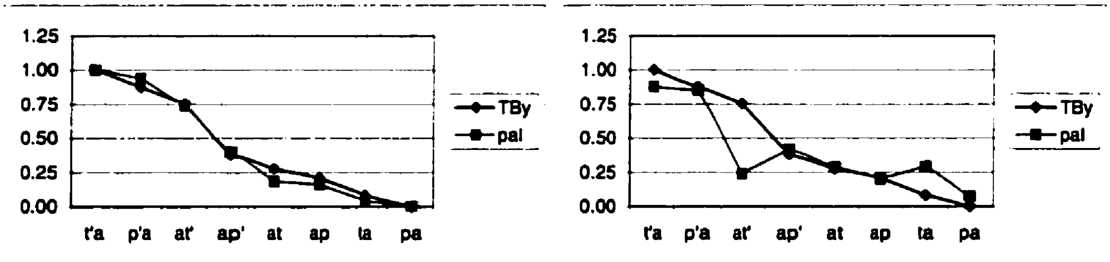


Figure 7.5 Correlation between Tongue Body (TBx) magnitude (max = 1, min = 0) and its perception (as palatalized = 1, as plain = 0), single labial consonants, condition RCI

The close correlation between TBx and the perceived plain-palatalized contrast in the onset consonants in clusters is very similar to the single initial consonants, and I do not examine it in detail. It is important to note that the absence of VC transition does not even slightly affect the perception of C_2, thus providing additional support for the primary importance of the CV transition and burst information.

As in the case of TBy, the perception of segments in terms of their secondary articulation in preconsonantal position is lower than would be expected based on TBx values in coda. This is shown in Figure 7.6. In general both /p'/ and /p/ are perceived as less palatalized than the corresponding degrees of Tongue Body fronting. The clusters /p’p’/, /p’p/, and /pp’/ are exceptions, apparently due to the higher degree of Tongue Body raising in these contexts (see Figure 7.1a).

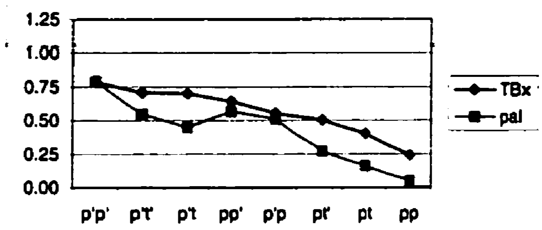


Figure 7.6 Correlation between Tongue Body (TBx) magnitude (max = 1, min = 0) and its perception (as palatalized = 1, as plain = 0), coda labial consonants (C_1) in clusters, condition RCI

In general, the correlation between the horizontal movement of Tongue Body and its perception in labials is modified by the environments in the same way as the vertical position of Tongue Body and its identification is. The distinction between coronals /t/ and /t'/, however, is more blurred.

How is the same articulatory input perceived by Japanese listeners (JCI) and by native listeners with noise (RNI)? Figure 7.7 presents the TBx values for the perception of single labials under the conditions JCI (a) and RNI (b). As we see, the two conditions show a very close interpretation of the TBx movement, in fact, even closer than the perception under RCI (see Figure 7.5). The latter was apparently influenced by some information from the VC transition and burst, while the Japanese responses and native responses with noise show a higher reliance on CV1.

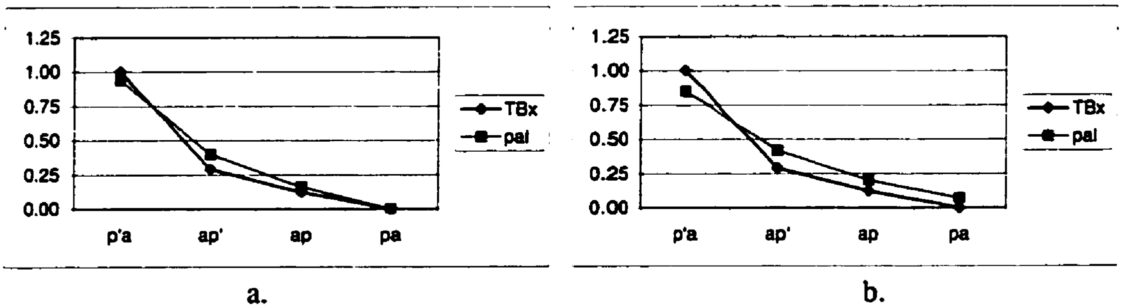


Figure 7.7 Correlation between Tongue Body (TBx) magnitude (max = 1, min = 0) and its perception (as palatalized = 1, as plain = 0), single consonants, conditions JCI (a) and RNI (b)

The perception of the plain/palatalized distinction in the other contexts under the conditions JCI and RNI follows the same pattern as in RCI, while showing more errors in the preconsonantal position. Also, the onset consonants in clusters in JCI (only /t/) and RNI (both /t/ and /p/) show an interesting effect of the preceding palatalized segment: in this environment they show a tendency to be perceived as palatalized. Thus, the Tongue Body gesture/feature is recovered but is attributed to a more preferred position for palatalization (onset) and often to a more likely candidate (/t/).

The written task provided interesting evidence that while recovery of the magnitude of the Tongue Body gesture/feature is relatively easy (both for native and non-native speakers), determining its timing with other gestures/features is a harder task. This is why both Russian and Japanese listeners showed a range of association patterns of the Tongue Body gesture/feature with consonants and vowels (mostly preceding). Based on these findings, it is not very surprising that the 'spreading' of the Tongue Body gesture/feature occurs in the less articulatorily and perceptually informative VC transition but not in the more informative CV transition and often affected the plain coronal /t/ (as C_1) (but not /p/ as C4), which is more susceptible to palatalization.

Determining the relative duration of the Tongue Body gesture/feature (one or two gestures/features) in the written task was also problematic, especially to the non-native speakers' unfamiliarity with the gestural properties of the stimuli.

7.2.3.3 Conclusion

In sum, a strong correlation exists between the TBx values and the perception of plain and palatalized stops under all the conditions. The perception of labials, /p/ and /p'/, shows a particularly high degree of dependence on the relative TBx magnitude. Listeners exhibited a preference for the information available at the CV transition. This is, in fact, where the initial /p/ and /p'/ are maximally differentiated by TBx.

In general, the predictions about the difference in perception of four stops based on their degree of Tongue Body raising/lowering (7.1) and fronting/backing (7.2) were fully confirmed. Recall these predictions: the palatalized gesture is primarily associated with /t'/, followed by /p'/. Further, it may be attributed to /t/, and least likely, to /p/.

(7.1) Degree of Tongue Body raising and perception of palatalization

t' > p' > t , p

(7.2) Degree of Tongue Body fronting raising and perception of palatalization

t', p', t > p

The articulatory asymmetries between the stops in Tongue Body timing and reduction (/p'/ vs. /t'/), as well as between the same stop in the initial and final or preconsonantal position (/p'/) were also reflected in perception. In fact, these asymmetries were often more substantial than would be expected solely on the basis of articulatory differences. This may be explained by perceptual and/or auditory factors: the reliance of listeners on CV transitions, often at the expense of the VC transition (and burst in the final environment) and the higher level of degradation of VC information with noise and under other conditions.

7.2.4 Lip Aperture and Tongue Tip

Based on the articulatory results in Chapter 5, I predicted that primary articulation is likely to play a minor role, if any, in the perception of the plain-palatalized distinction. In this section I examine the correlation between the categorical presence or absence of Lip Aperture or Tongue Tip and the reported perception of consonants as plain or palatalized. I also look at the Tongue Tip magnitude differences between /t/ and /t'/ found for Subject 3 (and Subject 2) and their relation to the perception of these consonants.

7.2.4.1 Results

Table 7.6 presents the results of the correlation between the primary articulator and responses 'palatalized' or 'plain'. A non-significant correlation of [rs = .13, p =.431] for RCI means that the recovered Tongue Tip gesture/feature did not necessarily correlate with whether a consonant was perceived as palatalized. Correspondingly, the Lip Aperture gesture/feature was not treated as indicative of a plain consonant.

Table 7.6 Correlation between the primary gesture and its perception

Condition	Perceived	TT/LA	
		rs	*p*
RCI	palatalized	0.13	0.431
JCI	palatalized	0.12	0.466
RNI	palatalized	0.18	0.271

The correlation was not significant for palatalization under any of the conditions.

Table 7.7 presents the results of the correlation between Tongue Tip magnitude (vertical dimension, TTy) and its perception. Significant correlations of [rs = -.99, p =.005] for /t/ and of [rs =.99, p =.007] for RCI mean that a lower position of Tongue Tip corresponded to the response /t/ and the higher value correlated with /t'/.

Table 7.7 Correlation between Tongue Tip magnitude (TTy) the and its perception as /t/ or /t'/

Condition	Perceived as	Tongue Tip	
		rs (4)	*p*
RCI	t	**-0.99**	**0.005**
	t'	**0.99**	**0.007**
JCI	t	-0.85	0.154
	t'	**0.96**	**0.044**
RNI	t	-0.59	0.409
	t'	0.52	0.485

The correlation was significant for both /t/ (negatively [rs = -.99, p = .005]) and /t'/ (positively) in RCI responses [rs = .99, p = .007]. For the Japanese listeners (JCI) it was significant only for /t'/ responses [rs = .96 p = .044]. It was not significant for either of the segments in the noisy stimuli.

7.2.4.2 Discussion

The finding that the primary articulators and perception of palatalization are not directly related is not surprising, since both labials and coronals can be either plain or palatalized. As we will see in section 7.3, the perception of place is dependent on the main correlate of palatalization, Tongue Body magnitude.

We also found that the relative height of Tongue Tip (TTy) corresponds to the responses of /t/ or /t'/ under two conditions. Thus, listeners may be sensitive to the minor articulatory differences in Tongue Tip magnitude. These differences are harder to recover with noise. It should be noted, however, that this result may be due to the more substantial Tongue Body differences rather than to Tongue Tip parameters. The current experiment does not make it possible to determine the relative contribution of the two articulators to the perception of /t/ and /t'/.

7.2.4.3 Summary

As expected, the perception of palatalization depends primarily on Tongue Body differences (TBy and TBx), reflected acoustically in the VC and CV transitions, as well as in burst properties. The higher and the more front position of Tongue Body, especially during the CV transition, was always interpreted as a correlate of a palatalized consonant, while a lower and more back Tongue Body was perceived as belonging to a plain segment. The direct application of this 'algorithm' by listeners, regardless of the consonant identity and specific environments, was the major source of errors related to palatalization.

7.3 Articulation and perception of place

How do the articulatory factors correlate with the perception of place of articulation? In this section I investigate the relation between primary and secondary articulators and the perception of stops as coronal or labial.

7.3.1 Lip Aperture and Tongue Tip

I begin with the perception of the primary place of articulation, Lip Aperture and Tongue Tip. Recall that these are, unsurprisingly, the main correlates of place: /p/ and /p'/ share the movement of Lips (LA), while coronals are characterized by the constriction of Tongue Tip at the upper teeth (alveolar ridge). Here I examine the correlation between the categorical presence or absence of Lip Aperture or Tongue Tip and the reported perception of consonants as coronal or labial.

7.3.1.1 Results

Table 7.8 presents the results of the correlation between the primary articulator and responses 'coronal' or 'labial'. A significant correlation of [rs = .98, p =.000] for RCI means that the recovered Tongue Tip gesture/feature was consistently labeled by listeners as 'coronal'. Correspondingly, the Lip Aperture gesture/feature was treated as a correlate of the labial consonant (the coefficient values for response 'labial' are not listed, since they are the reverse of the 'coronal' values, i.e. [rs = -.98, p =.000] for RCI).

Table 7.8 Correlation between the primary gesture and its perception

Condition	Perceived as	TT/LA	
		rs (40)	*p*
RCI	coronal	**0.98**	**0.000**
JCI	coronal	**0.85**	**0.000**
RNI	coronal	0.26	0.101

The correlation was significant for coronals in the RCI [rs = .98, p = .000] and JCI [rs = .85, p = .000] responses. It was not significant in the RNI condition.

7.3.1.2 Discussion

It was not surprising that the presence of Lip Aperture and Tongue Tip is identified with labial and coronal consonants respectively. These gestures/features are recovered based on VC and CV transitions and burst quality. What is more interesting is the fact that this does not hold of perception with noise. Obviously, it is harder to recover the place information under this condition due to the masked burst and blurred finer transition distinctions. As we will see further, listeners compensate by using Tongue Body differences to recover the place of consonants.

Before proceeding, it is useful to remind the reader about the acoustic consequences of the combination of primary and secondary articulators. The articulatory differences between plain labials (LA) and coronals (TT) are recovered from the spectral pattern (VC and CV transitions), where labials are typically associated with lower and coronals with higher F2 values. Similarly, the burst patterns of /p/ and /t/ have noise peaks at correspondingly lower and higher values. The addition of secondary palatal articulation, characterized by Tongue Body fronting and raising, substantially affects the spectral patterns of these consonants. F2 of both /p'/ and /t'/ is higher than that of the corresponding plain stops (recall, however, the timing differences; section 5.3.2.6.2). Their burst patterns also have energy at higher frequencies than for /p/ and /t/. In addition, the Tongue Body retraction during /p/ is likely to further lower the F2 formant for this segment. The fronting movement of Tongue Body due to Tongue Tip movement of /t/ also has its spectral consequences. As a result of all this, the maximum spectral contrast (and also the maximum Tongue Body difference) is between /p/ and /t'/, while the values for /t/ and /p'/ are in between on the spectral continuum. The acoustic pattern of /p'/ is the least homogeneous: since the Tongue Body peak is timed with the release of Lip Aperture, the consonantal spectrum shows high F2 at the CV transition and lower values for the same formant during the VC transition (however, higher than for /p/). In less informative environments (without CV or VC transition) and with additional noise this segment is the first candidate to be confused either with /t'/ (in the onset position) or with /p/ or /t/ (in the coda position).

These overall correlation results do not take into consideration the differences between environments. Below we will see how the perception of place varies in these contexts.

Unlike the discussion of Tongue Body magnitude (sections 7.2.2 and 7.2.3) where gradient articulatory values were used, here I use the categorical presence or absence of Lip Aperture and

Tongue Tip gestures. I use 1.00 to indicate the Tongue Tip gesture and .00 to denote the Lip Aperture gesture. The perception results may range from 1.00 (all responses are 'coronal') to .00 (identified as 'labial' 100% of the time). Thus, if perception were perfect, all the responses for coronals would have value 1.00 and those for labials .00.

Figure 7.8 plots the primary articulatory gestures (TT and LA) and the corresponding perception in single consonants. The value 1.00 for /ta/ both in terms of articulation and perception means that the initial /t/ was produced by the Tongue Tip gesture and was perceived as coronal 100% of the time. Looking at the results for /pa/, we can see that the initial plain labial /p/ was articulated by Lips (0.00) and was perceived as labial in 100% of the cases. Notice that while the Tongue Tip movement is consistently identified with coronals, there is somewhat more variation in the perception of final labials, particularly /p'/. Recall that the articulation of this consonant results in a spectral pattern that is closer to a coronal stop.

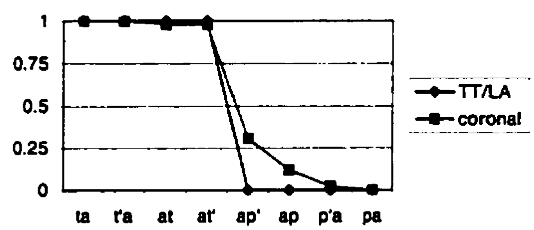
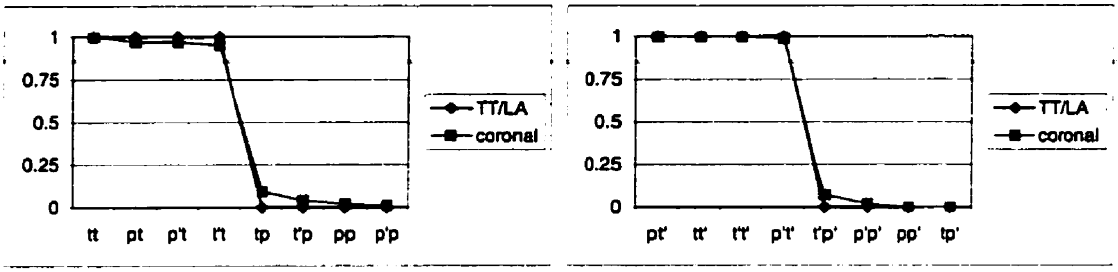


Figure 7.8 Correlation between production of primary gesture (TT = 1, LA = 0) and its perception (as coronal = 1, as labial = 0), single consonants, condition RCI

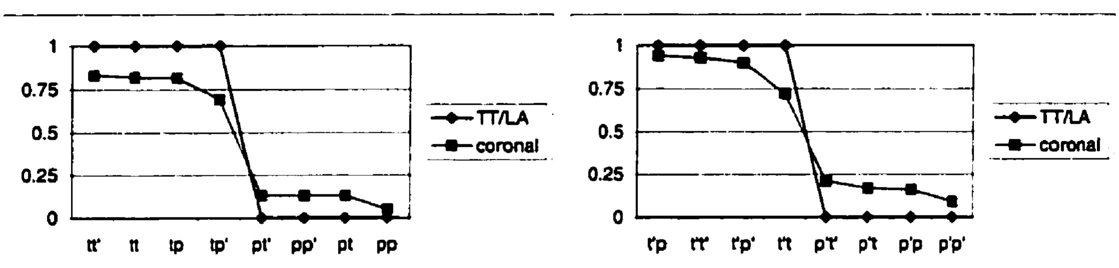
The burst information and CV transition for /t/ and /p/ appear to be sufficient to perceive the place contrast for the onset consonant in clusters even though the VC transition information is missing. Figure 7.9 shows the correlations between the primary articulators and their perception separately for labials (a) and coronals (b) (as C_2). We can see that the responses 'coronal' or 'labial' are highly related to the corresponding Tongue Tip and Lip Aperture gestures.



a. b.

Figure 7.9 Correlation between production of primary gesture (TT = 1, LA = 0) and its perception (as coronal = 1, as labial = 0), onset consonants (C_2) in clusters, plain (a) and palatalized (b), condition RCI

The perceived labial/coronal distinction is less categorical in coda consonants in clusters (Figure 7.10), where /t/ shows a slightly higher tendency to be perceived as labial than /t'/, and /p'/ is more often associated with coronals than /p/. Notice that /t'/ in the cluster /t't/ (Figure 7.10b) is confused with a labial more than 25% of the time. This is partly due to the lack of burst of /t'/ in this homorganic cluster (release rate is .00), which contains both place and palatalization information.

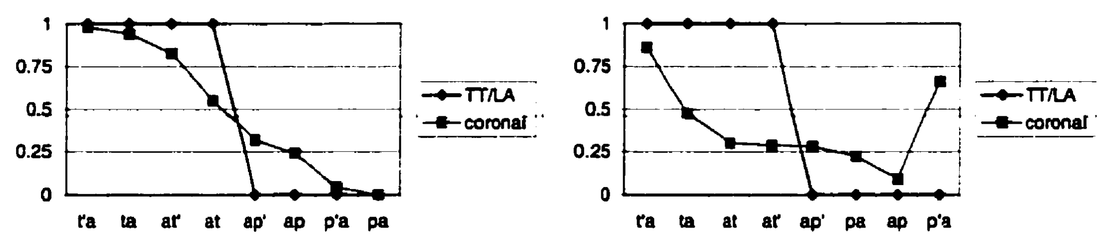


a. b.

Figure 7.10 Correlation between production of primary gesture (TT = 1, LA = 0) and its perception (as coronal = 1, as labial = 0), coda consonants (C_1) in clusters, plain (a) and palatalized (b), condition RCI

In general, the contrast in terms of place (in RCI) is more robust than the plain/palatalized distinction in all of the environments (see Figure 7.0-7.3). This can be explained by the fact that two different primary articulators are involved in the place distinction. These are fairly effectively differentiated both by transitions and burst. The presence of the secondary palatal articulation may enhance the contrast in some cases (e.g. /p/ vs. /t'/) but may also blur it in others (e.g. /p'/ and /t'/ or /p/ and /t/).

This can be seen in the perception of single stops by Japanese listeners and by Russian subjects with noise (Figure 7.11). While the Japanese listeners (a) have no problem recovering the place of initial stops, they often fail to do so with the final segments. In these cases, /t/ is more likely to be perceived as labial than /t'/, and /p'/ is a more likely candidate for coronal than /p/. Thus both the burst and transition information about the place of final stops is less recoverable than those of initial segments. The native speakers performing with noise (Figure 7.11b) take the differences between plain and palatalized segments of the same place of articulation and between positions even further apart. Only initial /t'/ and /p/ are closely matched with their primary articulators. The initial /p'/ is perceived as /t'/ in most of the cases. Recall that these two segments have almost the same CV transition and differ in burst, which is masked by the noise. Apparently, listeners do not attend to the differences between these segments at the VC transition. The final coronals show a very strong bias towards labiality. In the absence of most of the place information, listeners use the Tongue Body magnitude and slope during the CV transition, which is not always a reliable source of information for the final consonants.



a. b.

Figure 7.11 Correlation between production of primary gesture (TT = 1, LA = 0) and its perception (as coronal = 1, as labial = 0), single consonants, conditions JCI (a) and RNI (b)

7.3.1.3 Conclusion

In general, the perception of place in all environments correlates closely with the corresponding primary articulators. Initial stops are better recovered than final stops. There are asymmetries between labials and coronals in the degree of 'labiality' and 'coronality'. These problems are particularly exacerbated with the addition of noise.

7.3.2 Presence or absence of Tongue Body

We have seen that place perception is closely related to the plain/palatalized distinction of the stops. Does the categorical presence or absence of Tongue Body raising/fronting inform listeners about place of articulation? This correlation between Tongue Body gesture and the perception of consonants as coronal or labial is examined below.

7.3.2.1 Results

Table 7.9 presents the results of the correlation between the Tongue Body gesture and its perception. A significant correlation of rs = .41, p =.008 for RNI means that the presence of the Tongue Body raising/fronting gesture is related to the perception of this segment as coronal. The absence of this gesture is perceived by listeners as a correlate of a labial consonant (the correlation values for responses 'labial' are not listed, since the values for them are the opposite of the ones for 'coronal', i.e. the correlation between the presence of Tongue Body and response 'labial' in RNI is rs = -.41, p = .008).

Table 7.9 Correlation between the Tongue Body gesture and its perception

Condition	Perceived	TB fronting/raising gesture	
		rs (40)	*p*
RCI	coronal	0.04	0.785
JCI	coronal	0.07	0.649
RNI	coronal	**0.41**	**0.008**

This correlation was not significant for place of articulation under the conditions RCI [rs = .04, p = .785] and JCI [rs = .07, p = .649]. It was, however, significant for place (coronals) in RNI [rs =.41, p = .008].

7.3.2.2 Discussion

The correlation results show that both Russian and Japanese listeners, presented with clean stimuli, do not associate Tongue Body raising/fronting with a particular place of articulation. However, the perception of the noisy stimuli by native speakers shows significant results for this correlation. Apparently, noise makes it difficult to differentiate between the movements of Tongue Tip and Tongue Body (based on the VC and CV transitions). The detected fronting and raising of Tongue Body can be interpreted as the Tongue Tip fronting/raising gesture (which also involves Tongue Body fronting), and vice versa. Obviously, this can lead to confusion between palatalized consonants (/p'/ and /t'/) and coronals (/t/ and /t'/). This finding is of particular interest, since it suggests that the common process of change of palatalized consonants of any place to coronal place of articulation (Bhat 1978, Hume 1992), as well as the neutralization patterns of /t/ and /t'/ (in either direction) are motivated by recoverability of (and confusion between) the Tongue Tip and Tongue Body gestures/features.

In the next sections I examine this correlation in more detail, looking at the vertical and horizontal dimensions of Tongue Body at several points in time.

7.3.3 Tongue Body

The correlations discussed in this section consider whether listeners use the relative Tongue Body magnitude (both vertical, TBy, and horizontal, TBx, dimensions) at four points in time (VC1, VC2, CV1, and CV2) and Tongue Body slope (during the VC and CV transitions) to determine the coronal/labial distinction.

Recall that while the vertical Tongue Body movement differentiates the plain-palatalized distinction for both labials and coronals, the horizontal movement of this articulator serves to differentiate the same contrast only between plain and palatalized labials. The Tongue Body during the production of /t/ is also fronted. This differentiates the plain /p/ from all other stops (/t/, /p'/, and /t'/).

7.3.3.1 Results

Tables 7.10 and 7.11 present the results of the correlations of Tongue Body magnitude. A significant correlation of rs = .74, p =.000 for RNI at CV1 means that higher values of TBy at this point were consistently associated with the coronal place of articulation of the segment. The lower values were interpreted as denoting a labial consonant.

Table 7.10 Correlation between the Tongue Body (y) raising and its perception at four points in time

Condition	Perceived	TBy VC1		TBy VC2		TBy CV1		TBy CV2	
		rs	*p*	*rs*	*p*	*rs*	*p*	*rs*	*p*
RCI	coronal	0.14	0.506	0.26	0.215	0.14	0.501	0.27	0.207
JCI	coronal	0.22	0.229	0.31	0.137	0.15	0.481	0.27	0.198
RNI	coronal	0.06	0.773	0.24	0.250	**0.74**	**0.000**	**0.73**	**0.000**

Table 7.11 Correlation between the Tongue Body (x) fronting and its perception at four points in time

Condition	Perceived	TBx VC1		TBx VC2		TBx CV1		TBx CV2	
		rs	*p*	*rs*	*p*	*rs*	*p*	*rs*	*p*
RCI	coronal	0.07	0.733	**-0.56**	**0.005**	-0.28	0.191	-0.05	0.815
JCI	coronal	0.15	0.488	**-0.45**	**0.029**	-0.29	0.195	-0.05	0.835
RNI	coronal	-0.02	0.946	-0.29	0.167	**-0.80**	**0.000**	**-0.78**	**0.000**

The correlation between perceived place and TBy magnitude was found to be non-significant at all four points in time for the conditions RCI and JCI. But this correlation was significant for RNI during the CV transition (both CV1 and CV2). There was a significant correlation between place and TBx magnitude at VC2 in RCI and JCI. Under RNI the place was significantly correlated with TBx at CV1 and CV2.

Tables 7.12 and 7.13 show the results of the correlation of Tongue Body slope and perceived place. A significant correlation of rs = -.72, p =.000 for RNI (response 'coronal') during the VC transition means that a lowering movement (a negative slope) of Tongue Body was associated by listeners with noise with the coronal place of articulation of the segment.

Table 7.12 Correlation between the Tongue Body (y) slope during VC and CV transition and its perception

Condition	Perceived as	TBy VC slope		TBy CV slope	
		rs	*p*	*rs*	*p*
RCI	coronal	0.35	0.095	-0.01	0.960
JCI	coronal	0.37	0.078	-0.02	0.929
RNI	coronal	0.39	0.060	**-0.72**	**0.000**

Table 7.13 Correlation between the Tongue Body (x) slope during VC and CV transition and its perception

Condition	Perceived as	TBx VC slope		TBx CV slope	
		rs	*p*	*rs*	*p*
RCI	coronal	**0.71**	**0.000**	**-0.63**	**0.000**
JCI	coronal	**0.61**	**0.002**	**-0.63**	**0.001**
RNI	coronal	0.35	0.094	**-0.73**	**0.000**

There was a correlation between perceived place and CV slope of the vertical Tongue Body only for RNI. Place was significantly correlated with the horizontal Tongue Body slope (both the VC and CV slope for RCI and JCI; only CV slope for RNI).

7.3.3.2 Discussion

These results confirm that TBy and TBx information is used differently in the identification of place, as well as that noise affects the choice of relevant information. While only TBx is normally necessary for the perception of place, both TBx and TBy can be taken into account in the lack of other place information sources. This results in higher confusion between plain/palatalized (particularly /t/ and /t'/) and labial/coronal (especially /p'/ and /t'/) consonants with noise.

Note also that both Russian and Japanese listeners, consult the TBx magnitude at VC2 (together with the VC and CV TBx slope) when determining the place of articulation under the clean condition. This is where the place differences are at their maximum (see 5.3.2.6.2). Since the VC information is more susceptible to noise, the listeners under this condition are left with CV differences, which are relatively minor.

7.3.3.3 Conclusion

In sum, Tongue Body information, being an important correlate of the plain/palatalized contrast, is used to a limited extent in the perception of the place distinction. Adding noise masks important place information. In order to recover this distinction, listeners use Tongue Body movement as an indicator of both place and palatalization contrasts. This, as we predicted, increases the confusion between labials and coronals on the one hand, and plain and palatalized segments on the other.

7.4 Articulation and correct identification

7.4.1 Tongue Body difference

In previous sections we examined the correlations between relative Tongue Body magnitude values and their association with the *perceived* plain/palatalized contrast. Not all of those responses were correct. As we saw, this was partly due to the variation in Tongue Body values in various environments. In this section I examine the role of the overall Tongue Body differences (at all points) in various contexts in the *correct* identification of segments. The question is whether the overall amount of articulatory difference (section 5.3.4) in a given environment correlates with how correctly the stops are identified.

7.4.1.1 Results

Table 7.14 presents the results of the correlation. The significant results are given in bold. A significant correlation of [$r_s = .65$, $p = .000$] for RCI means that the more the overall difference between TBy values in a given context, the higher the correct identification of consonants. The smaller difference correlates with lower recognition rate and more confusion of segments.

Table 7.14 Correlation between the overall Tongue Body (TBy and TBx) differences and correct identification

Condition	Correct Ident	TBy difference		TBx difference	
		rs	*p*	*rs*	*p*
RCI	total	**0.65**	**0.000**	**0.71**	**0.000**
JCI	total	**0.32**	**0.045**	**0.34**	**0.034**
RNI	total	**0.70**	**0.000**	**0.79**	**0.000**

A significant correlation between the TBy and TBx differences and correct identification of magnitude was found for all three conditions.

7.4.1.2 Discussion

The correlation results show that the correct identification rate is dependent on the articulatory Tongue Body differences between consonants in a given environment. Recall that the segments were maximally differentiated by Tongue Body when single in the initial environment, and had minimum differences in the preconsonantal position.

In section 7.3 we determined that listeners do not attend to all articulatory differences to the same degree, showing more reliance on the information provided by the CV transition than by the VC transition. Thus, the substantial articulatory differences between single initial (onset) consonants and onset segments in clusters do not result in the same perceptual differences. In fact, their recognition is almost the same. Also, the relative articulatory differences between final single consonants and preconsonantal segments are not as great in perception because listeners evaluate single consonants based mostly on the CV transition (as if they were initial segments), ignoring the VC transition information. Note that here I used the total identification rate, regardless of place or palatalization.

7.4.1.3 Conclusion

Recall the predictions about overall Tongue Body difference for perception: less substantial articulatory differences between TB gestures would result in poorer recoverability. These predictions are generally supported. However, the results show that listeners are selective: the differences at the CV transition are more important than VC differences, even if they are of the same magnitude. Thus, while the statements (7.3acd) appear to be true both in terms of articulation and perception, statement (7.3b) is only partly supported by the identification results. While initial and final single consonants are generally better identified than the same segments preconsonantally, the single consonants are not necessarily better recognized than the onset segments in clusters. In fact, single final consonants repeatedly show lower identification rates than prevocalic consonants of a cluster.

(7.3) Overall Tongue Body differences and perception

a. C_i vs. C_j/(V)#_V > C_i vs. C_j/V_#(V)

b. C_i vs. C_j/(V)#_V or V_#(V) > C_i vs. C_j/(V)_C or C_(V)

c. C_i vs. C_j/ C_(V) > C_i vs. C_j/(V)_C

d. C_i vs. C_j/_C > C_i vs. C_j/_C'

In sum, the overall Tongue Body differences affect the correct identification of consonants, but the identification results also shows that listeners attend to some information and ignore other differences.

7.4.2 Release rate and lag

The analysis of presence or absence of stop burst allows us to predict that consonants will be better recovered when they are audibly released (section 5.5), before vowels and before other consonants, provided there is a positive lag between the two constrictions. The latter is more common for hetero-organic back-to-front clusters (/tp/, /t'p/, /tp'/, and /t'p'/. We also emphasized the differences between stops in burst quality: palatalized stops had a more salient burst than the plain ones and coronals showed a higher energy burst than labials (section 5.5.1).

Below I investigate the correlations between these articulatory and acoustic factors and the correct identification of the consonants. It is not easy to test burst quality in the absence of actual acoustic measurements. The measure I used (burst quality/release rate) assigns relative degree of salience to stop bursts (/p/ = 1, /p'/ = 2, /t/ = 2, and /t'/ = 3) and multiplies it by the release rate of a given stop in a given environment. Thus, the initial /p/ has a value of 1.00 (burst quality 1.00 * release rate 1.00), while the initial /t'/ is assigned 3.00 (3.00 * 1.00). Along the same lines, /p/ in the cluster /pt/, with a release rate of .2, is given the value of .2 (1.00 *.2), while /t'/ in the sequence /t'p/, with a release rate of 1.00, would still receive 3.00 (3.00 * 1.00).

The correlation between lag and correct identification is based on a small number of number of values, 8.

The task of identification with noise (RNI) provides an additional test for the importance of these factors, since under this condition burst is masked, and thus release rate, burst quality, and lag should not be of any importance.

7.4.2.1 Results

Table 7.15 presents the results of the three correlations. A significant correlation of [rs = .71, p =.000] for RCI means that the more often the stops were released, the higher was the correct identification rate. Correspondingly, lower release rate corresponded to more confusion.

Table 7.15 Correlation between the correct identification and release rate, burst quality/release rate, and lag

Condition	Correct Ident	Release rate		Burst quality/ Release rate		Lag	
		rs	*p*	*rs*	*p*	*rs*	*p*
RCI	total	**0.71**	**0.000**	**0.61**	**0.000**	0.60	0.113
JCI	total	**0.63**	**0.000**	**0.45**	**0.003**	-0.01	0.976
RNI	total	0.23	0.150	0.17	0.286	-0.26	0.540

The correlation between the correct identification and release rate was significant for RCI and JCI, but not for RNI. Similarly, there was a significant correlation between correct identification and burst quality/release rate under the same conditions. The correlation between correct identification and lag was not significant under all the conditions. However, it had a marginal value [rs = .70, p = .055] for palatalized consonants in RCI. The low coefficient for lag was based only on 8 values.

7.4.2.2 Discussion

The results show that correct identification of consonants is dependent on the release rate and the quality of burst for both Russian (clean) and Japanese listeners. Released stops are easier to identify, while no release leads to more perceptual confusion. Also, a more salient burst (together with the release rate) corresponds to a higher identification rate of stops. Adding noise (RNI) blurs the perceptual difference in terms of release rate and burst quality, leading to an overall lower recognition rate.

It should be noted that the current analysis does not allow one to distinguish the relative contribution of the release rate from the Tongue Body differences in single consonants and onset consonants in clusters.

The influence of articulatory lag on identification was marginally significant in RCI. This could be due to the overall small number of clusters that exhibited positive lag.

Recall that the results of the writing task show a very high degree of missed consonants as C_1 in clusters, especially by the Japanese listeners. Although no correlation between this fact and release rate was made, they are likely to be related. The lack of burst and higher overlap with the following consonant has been found to result in the perception of clusters as single consonants with properties of C_2 rather than C_1 (Ohala 1990). The same factor, together with the poor VC transitions, is also likely to have contributed to the wide range of responses for C_1 in clusters in the written task.

7.4.2.3 Conclusion

The results show the importance of burst release rate and quality, confirming our predictions about perception based on articulatory and acoustic differences. Higher release rate (7.4 and 7.5) and more salient burst (7.6) result in better identification of the corresponding stops. Based on the current perceptual experiment and correlations, however, it is not easy to determine the relative contribution of each of these factors separately from each other and from other articulatory differences. For instance, it is obvious that the perceptibly poor burst of /p/ does not affect the overall high identification of this stop, which is well differentiated from other stops by the combination of articulatory movements of Lip Aperture, TBy and TBx (see also Byrd 1992 about asymmetries in the recoverability of labials and coronals).

(7.4) Presence/absence of audible release depending on position in cluster and identification
C/C_V > C/V_C

(7.5) Presence/absence of audible release depending on the place of the following consonant and identification
C_i/_C_j > C_i/_C_i

(7.6) Acoustic burst quality and perception
a. /t'/, /t/ > /p/, /p'/
b. /t'/ > /t/
c. /p'/ > /p/

7.5 General conclusion

A strong dependence exists between the articulatory, and the resulting acoustic, differences between consonants and the recoverability of plain and palatalized labial and coronal stops. Relative Tongue Body magnitude and slope, primary gestures, release rate and burst quality show significant correlations with the perception and identification of the four consonants under study. In addition, the analysis helped us identify additional factors that affect the perception of the palatalization and place contrasts, namely listeners' general bias towards CV transitions often at the expense of VC transitions.

In the next chapter I show how these articulatory, acoustic, and perceptual factors are instrumental in deriving the observed phonotactic patterns. In order to do that I use the experimental results to create a perceptual scale denoting how likely a given segment in a given environment is to be perceived as palatalized. The scale (Table 7.16) is based on three conditions, RCI (Russian, clean, identification), RNI (Russian, noise, identification), and JCI (Japanese, clean, identification). The items (stimuli) are ordered from the highest value at the top (most likely to be perceived as palatalized) to the lowest value (least likely to be identified as palatalized).

Table 7.16 Perceptibility scale: Perceived as palatalized, combined for conditions RCI, RNI, and JCI

a. Single consonants

	single
ta **t'apy**	0.96
ta **p'apy**	0.93
tat' apy	0.61
tap' apy	0.46
tat apy	0.16
tap apy	0.15
ta **tapy**	0.12
ta **papy**	0.03

b. Consonants in clusters

	onset		coda
tat' **t'apy**	0.95	**tat'** p'apy	0.67
tat **t'apy**	0.95	**tat'** t'apy	0.63
tap **t'apy**	0.94	**tap'** p'apy	0.60
tat' **p'apy**	0.93	**tat'** papy	0.58
tat **p'apy**	0.92	**tat** t'apy	0.57
tap' **t'apy**	0.92	**tat'** tapy	0.52
tap **p'apy**	0.91	**tap** p'apy	0.48
tap' **p'apy**	0.86	**tap'** t'apy	0.45
tap' **tapy**	0.36	**tat** p'apy	0.43
tat' **tapy**	0.24	**tap'** tapy	0.39
tat **tapy**	0.13	**tap'** papy	0.34
tat' **papy**	0.12	**tat** papy	0.32
tat **papy**	0.10	**tap** t'apy	0.31
tap **papy**	0.09	**tat** tapy	0.22
tap' **papy**	0.08	**tap** papy	0.21
tap **tapy**	0.04	**tap** tapy	0.16

I assume that this scale, a combination of perceptual results under different conditions, neutralizes some language-particular effects and differences in experimental conditions, and thus gives us an approximation of recoverability in general. The scale will serve as an input to our hypothetical language learner in Chapter 8.

Chapter 8. Deriving the patterns

8.0 Introduction

This chapter reviews the phonetic factors identified in the previous chapters and uncovers the crucial role they play in the emergence of cross-linguistically common palatalization patterns.

8.1 The model of neutralization: Review

Recall from Chapter 1 that I assume a model that accounts for neutralization of a phonological contrast, either active or static, which involves articulation, acoustics, and perception on the one hand, and cognitive processing of the recovered information on the other hand. This model, repeated from Chapter 1, is outlined in Figure 8.1.

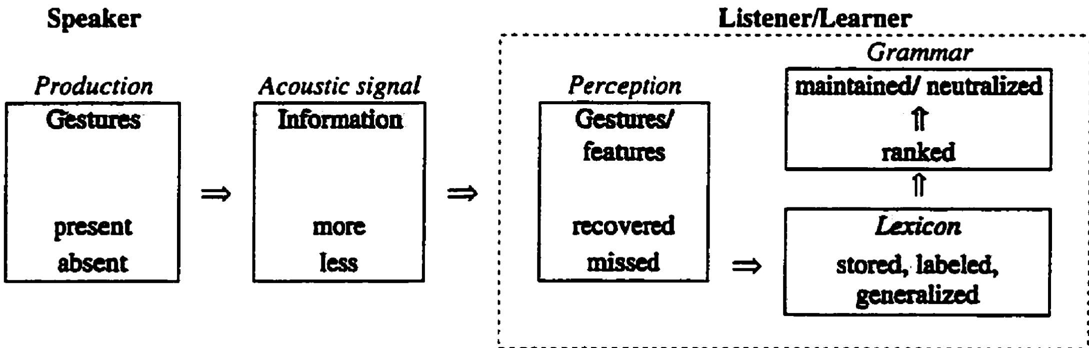


Figure 8.1 The model of neutralization of a segmental contrast

Consider an example. A speaker produces two utterances that have two segments which are contrastive in her/his language in some environment. The two segments differ articulatorily as to which gestures are involved in their production (articulatory differences). The output of articulation is an acoustic signal that contains information about the articulatory gestures/features (acoustic information). This information is extracted from the acoustic signal and interpreted by a listener/learner. I refer to this process as recoverability of gestures/features. Recall from Chapter 1 that I use the term gesture/feature to refer to the gesture as an abstract phonological unit, rather than a motor program. The recovered gestural/featural representation is further stored in memory and classified with reference to the available lexical and phonological categories. This is used as a basis for constructing, or modifying, constraints in the phonological grammar.

Both articulation and perception are affected by a number of factors. Some are external to the grammar (e.g. gestural reduction (Kelso. Saltzman, Tuller 1986, Gick 1999) or are due to failure to recover a gesture/feature (e.g. Surprenant & Goldstein 1998, Browman & Goldstein 1999)), or can be attributed to the lower, phonetic levels of this knowledge (e.g. interpretation of acoustic patterns). The recovery of gestures/features by the listener/learner is the key to our analysis. The more information about input gestures/features is available to the listener/learner, the more likely it is that s/he can recover them. Fewer robust articulatory and acoustic differences result in failure to recover these phonological units or in the confusion of them with other features/gestures.

Ease of recovery has direct consequences for the lexicon and the grammar. As a result of this, and additional cognitive factors, the phonological contrast in the grammar is either maintained or neutralized. The details of the process are discussed below with reference to overall results of the phonetic experiments presented in the earlier chapters.

8.2 The interaction of phonetic and cognitive factors

In this section I demonstrate how common phonotactic restrictions can arise from the interaction of phonetic factors with phonological structural constraints. This is done using a hypothetical learning situation, which, undoubtedly, is a very coarse model of a real-life acquisition path. Nevertheless, it provides us with an important insight into the mechanism of neutralization.

8.2.1 A hypothetical learner

Imagine the following scenario. A hypothetical language, Language X, has the following consonant inventory {p, p', t, t'} and a vowel {a}. The lexicon of the language consists of the following items (8.1): monomorphemic CV and VC stems (stems 1 and 2: 8.1ab), suffixed VCV forms (8.1c), and VCCV compounds of the type 'stem 2 + suffix' (8.1d).

(8.1) Lexicon of Language X

	Syllable structure	Lexical items	Morphological categories	Morpheme structure
a.	CV	*pa, p'a, ta, t'a*	Stems 1	*pa, p'a, ta, t'a*
b.	VC	*ap, ap', at, at'*	Stems 2	*ap, ap', at, at'*
c.	VCV	*apa, ap'a, ata, at'a*	Suffixed forms	Stem 2 + *-a;*
d.	VCCV	*appa, app'a, apta, apt'a* *ap'pa, ap'p'a, ap'ta, ap't'a* *atpa, atp'a, atta, att'a* *at'pa, at'p'a, at'ta, at't'a*	Compounds	Stem 2 + Stem 1

Assume also that our hypothetical learner has already acquired some knowledge of the phonemic categories /p/, /p'/, /t/, /t'/, but has not yet learned the constraints on their distribution (cf. Hayes 1999). The learner is exposed to the words in (8.1) in a random order. Each of these words in the same context is presented 100 times.

The goals of the learner are the following. First, when presented with a stimulus (e.g. *ap'a*), s/he recovers the gestures/features from the speech signal and stores this information in memory. I refer to these two tasks as *recoverability* and *storing*. Second, based on all stored instances, or tokens, of this item the learner determines a surface form for each lexical item (e.g. [ap'], [ap'a], etc.). A decision algorithm for this procedure, referred to as *labeling*, is discussed below. Third, the learner examines all lexical items that share the same morphemes (e.g. [ap'], [ap'a], [ap'ta], etc.) and posits underlying forms for each of them (e.g. /ap'/). I refer to this task as *generalizing*. Fourth, surface forms, underlying forms, and alternations between them are examined for distribution of segments (or gestures/features) in each environment. Based on this the learner constructs the grammar by stating generalizations on the occurrence of segments or gestures/features in various environments, the process referred to as *ranking*.

Assume that these generalizations take the form of Optimality Theory-type constraints (Prince & Smolensky 1993). The learner is provided with markedness constraints on segments or features/gestures and the faithfulness constraint Ident[feature] (8.2).[75] The markedness constraints prohibit segments in specific environments (8.2a). The faithfulness constraint (8.2b) which requires that an input feature is maintained in the output, is ranked against the markedness constraints.

(8.2) a. *p'/_V: Disallow /p'/ prevocalically;
*p'/_#: Disallow /p'/ word-finally;
*p'/_p: Disallow /p'/ before /p/;
*p'/_t: Disallow /p'/ before /t/;
*p'/_p': Disallow /p'/ before /p'/;
*p'/_t': Disallow /p'/ before /t'/;

b. Ident[pal]: Maintain the specification of gesture/feature with respect to palatalization ([+pal] or [-pal]) in the input.

It is important for our analysis that the markedness constraints are not preranked, i.e. the learner is not aware of the markedness scale. I assume that these unordered constraints on each segment (/p/, /t/, /p'/, /t'/) in each environment (_V, _#, etc.) are supplied by UG (8.3).

(8.3) a. *p/_V, *p/_#, *p/_p, *p/_t, *p/_p', *p/_t';
*p'/_V, *p'/_#, *p'/_p, *p'/_t, *p'/_p', *p'/_t';
*t/_V, *t/_#, *t/_p, *t/_t, *t/_p', *t/_t';
*t'/_V, *t'/_#, *t'/_p, *t'/_t, *t'/_p', *t'/_t';

b. Ident[pal].

The ranking of the faithfulness constraint against the markedness constraints, as well as the relative ranking of markedness constraints with respect to each other, are determined by the learner based on the lexicon (surface forms and underlying forms).

I also assume that the learner can make higher-level generalizations over the specific constraints to come up with constraints that refer to more abstract categories like segment classes (features/gestures, e.g. *[+pal]/_# referring to all palatalized consonants) and more general

[75] Note that the constraints on statements (e.g. *p'/_V) can be alternatively stated in terms of gestures/features

environments (word edges, onset or coda, e.g. *p']$_\sigma$ referring to the coda position, or both final and preconsonantal environments).

This grammar would allow the learner to generate and recognize all possible well-formed forms of the language (as exhibited by language speakers in various productive behaviours, such as new word coinage, loan adaptation, well-formedness judgements, etc.).

Crucial to this situation is that the learner's ability to recover gestures/features correctly from the acoustic signal is limited, a view that is commonly ignored in the learnability work in OT (e.g. Tesar & Smolensky 1993, 1998). A gesture/feature can be recovered, missed, or confused with another gesture/feature. This has important consequences for what the learner stores in memory, what underlying forms s/he posits and, ultimately, what ranking of constraints s/he arrives to.

To model the learner's perception I use the scale derived from our perceptual experiments. This scale is based on the combined results for perceived palatalization under three conditions (see Chapter 7). Clearly, this is a very rough approximation of a real learner's perception, nevertheless, it is not unreasonable. The limitations of a child's perception may not be entirely different from an adult's perception. In fact, the learner's perception may share some characteristics with different groups of subjects and conditions examined in our perception study. First, like the Japanese listeners in our perceptual experiments, the hypothetical learner can recognize the phonemic categories of the input, but is not yet familiar with the phonotactics of the presented utterances. At the same time the learner is somewhat closer to the Russian listeners in terms of familiarity with the details of the gestural input. The combination of the clean and noisy perception results for Russian listeners is likely to be closer to the real life acquisition situation, since language acquisition rarely happens in complete silence (or in headphones), nor does it occur in substantial noise.

---

*[lab][pal][-cont] (or [Lips-closed][TB-narrow]). These notational differences are not crucial for our analysis.

The perceptibility scale is given in Table 8.1. For the sake of simplicity I ignore word boundaries and other consonants and vowels that were present in the stimuli used for the scale (see Chapter 7).

Table 8.1 Perceptibility scale: Perceived as palatalized, combined for conditions RCI, RNI, and JCI (repeated from Chapter 7)

a. Single consonants

t'a	0.96
p'a	0.93
at'	0.61
ap'	0.46
at	0.16
ap	0.15
ta	0.12
pa	0.03

b. Consonants in clusters

	onset		coda
att'a	0.95	at'p'a	0.67
at't'a	0.95	at't'a	0.63
apt'a	0.94	ap'p'a	0.60
at'p'a	0.93	at'pa	0.58
ap't'a	0.92	att'a	0.57
atp'a	0.92	at'ta	0.52
app'a	0.91	app'a	0.48
ap'p'a	0.86	ap't'a	0.45
ap'ta	0.36	atp'a	0.43
at'ta	0.24	ap'ta	0.39
atta	0.13	ap'pa	0.34
at'pa	0.12	atpa	0.32
atpa	0.10	apt'a	0.31
appa	0.09	atta	0.22
ap'pa	0.08	appa	0.21
apta	0.04	apta	0.16

A scale of perceived palatalization indicates how often responses for a given token are perceived as palatalized or plain. .00 stands for the response 'plain' and 1.00 denotes the answer 'palatalized'. For our purposes, this means that in the former case (.00) none of 100 tokens of a lexical item are perceived as having a palatalized consonant. In the latter case (1.00) all 100 instances are perceived with a 'palatalized' segment. Note that such perfect perception is never the case. For example, when presented with 100 tokens of the item *p'a*, the learner identifies 93 of them as [p'a] and 7 as [pa] (the score of .93). The other items may show more variation. Out of 100 tokens of *ap'* only 46 are correctly identified as [ap'] (the score of .46). Other 54 instances are perceived as [ap].

The degree of variation in perception is important for labeling, the task that involves assigning surface forms to each lexical item. While in the case of *p'a* the learner would have little doubt about the surface form [p'a] rather than [pa], s/he is forced to make a decision between almost equally likely candidates [ap'] and [ap].

For the sake of exposition I consider four possible lexicons (surface forms) that differ in how much variability they allow, as shown in Table 8.2. If a consonant in a lexical item in Lexicon 1 has a score (the mean for all the stored exemplars for this item) of less than .45, then it is labeled as 'plain'. If it is higher than .55 it is considered 'palatalized'. If the item falls in the range .45-.55 (i.e. up to .05 from the mean value, .50), it is randomly assigned either of the two labels. Other lexicons are based on the same principles varying in the range of variation: .10 for Lexicon 2, .15 for Lexicon 3, and .20 for Lexicon 4. This will be crucial in determining the areas of the lexicon that are the least and the most stable and gradient differences between them.

Table 8.2 The hypothetical lexicons (surface forms) and their labeling procedures

	Labeled as		
	'plain'	'plain' or 'palatalized'	'palatalized'
Lexicon 1	.00-.45	.45-.55	.55-1.00
Lexicon 2	.00-.40	.40-.60	.60-1.00
Lexicon 3	.00-.35	.35-.65	.65-1.00
Lexicon 4	.00-.30	.30-.70	.70-1.00

I examine the 'learning' situation separately for single consonants, and onset and coda consonants in clusters. In each case I begin with a brief overview of phonetic factors, show how they restrict the choice of possible lexicons, and, together with additional factors, contribute to the selection of specific phonotactic grammars.

8.2.2 Constraints on single consonants

In this section the inputs to the learner consist of the words with single consonants in the initial onset (*pa, p'a, ta, t'a*) and final coda positions (*ap, ap', at, at'*).

8.2.2.1 Phonetic factors and perception

Recall that the articulatory differences in the Tongue Body gesture (fronting/backing and raising/lowering) which characterize plain and palatalized consonants are at their maximum in syllable onset, or prevocalic, position. In the final coda the palatalized gesture shows some reduction in magnitude, which is motivated by the dynamics of gestural movement. It may also undergo a shift in timing with the primary gesture. The acoustic consequence of these processes is that the CV and VC transitions and burst which are used by listeners to recover the stop gestures are less robust in final position (if not phrase final) than in prevocalic position.

In addition, the stops /p'/ and /t'/ vary in their realization of the palatalized gesture. Tongue Body is articulatorily more independent from the Lips in the production of /p'/ (more similar to [j]) and is more affected by reduction and shift. The palatalized gesture of /t'/, coupled with the Tongue Body, is additionally influenced by the fronting and raising of the primary articulator and thus is higher and, possibly, more front than the same gesture of /p'/. The resulting acoustics of the palatalized labial are different from the palatalized coronal in some important ways. First, the high second formant (F2) characteristic of Tongue Body fronting and raising is lowered for /p'/ due to the consequence of Lips movement, characterized by low F2. Lower F2 is more apparent during the VC transition to the onset /p'/ and during the CV transition of the final /p'/ (before a vowel), due to the shift of the Tongue Body gesture. Finally, the aero-dynamic factors that accompany the constrictions of /p'/ and /t'/ result in a more salient burst for the latter segment compared to the former. The burst release is always present in the onset and occurs optionally in final position (always before a vowel, often phrase finally, and less commonly before a consonant, depending on language-particular release strategies).

All this results in better or worse recoverability of a consonant in a given environment. Figure 8.2 shows the scales for plain and palatalized labials (a) and coronals (b) where these consonants are perceived by listeners as palatalized, based on the combined values for three conditions (from Table 8.1). 1.00 means that a segment is perceived as palatalized 100% of the time, and .00 for a 100% perception of the input as plain.

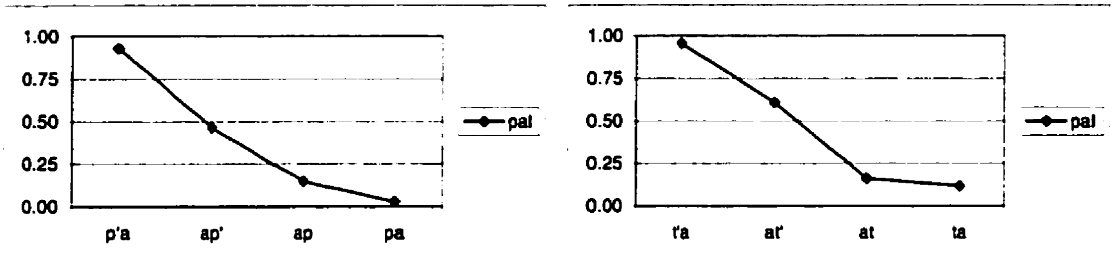


Figure 8.2 Perception of labials (a) and coronals (b) as palatalized, single consonants, combined means of conditions RCI, RNI, and JCI

Notice that the consonants which deviate most from their ideal recoverability are the palatalized /p'/ and /t'/ in final position, with the score of the labial /p'/ being slightly below the category boundary, .50, which means that it is perceived as plain as frequently as it is considered palatalized.

It is important that the perceptual scales in Figure 8.2 result from the 'blind' application of several independent factors. The listener/learner recovering the gestures/features from the signal does it without necessary awareness of all the fine-grained articulatory and acoustic factors involved. By doing this s/he either correctly recovers the gestures or fails to do so.

Below I show what consequences this may have for the lexicon constructed by our learner. I then examine a set of grammars that can be constructed on this basis.

8.2.2.2 Possible lexicons and grammars

In Table 8.3 the lexical items from the source lexicon (Lexicon 0) are ordered in terms of their perception as plain or palatalized, based on the scores given under 'Perceived as palatalized'. Recall that a score of .96 means that the consonant of the item in question was classified as 'palatalized' in 96 tokens and as 'plain' in 4 instances out of 100. The horizontal line marks the division between the two categories (.50). We can see that the word *t'a* is at the top, having the highest perception as palatalized (.96), and the item *pa* is at the bottom, perceived as plain almost all of the time (.03). The value for *ap'* is given in bold to show that although palatalized in the input, the perception of this item is below .50, and thus in the category 'plain'.

Table 8.3. Lexicons (surface forms) that can be derived based on the perceptibility scale for single consonants

Lexicon 0	Perceived as palatalized	Lexicon 1 *0.05*	Lexicon 2 *0.10*	Lexicon 3 *0.15*	Lexicon 4 *0.20*
t'a	0.96	t'a	t'a	t'a	t'a
p'a	0.93	p'a	p'a	p'a	p'a
at'	0.61	at'	at'		
ap'	**0.46**				
at	0.16	at	at	at	at
ap	0.15	ap	ap	ap	ap
ta	0.12	ta	ta	ta	ta
pa	0.03	pa	pa	pa	pa

The columns to the right show how these lexical items may be labeled in the four lexicons based on the procedures defined above. The ambiguous perceptual status of *ap'* leads to representing it as either plain or palatalized in all the lexicons, as shown by shading. The word *at'* with the final /t'/ is the next confusing item. It is stored either as /at'/ or /at/ in Lexicons 3 and 4. All the other words are relatively far from the variation range and thus are unlikely to be mislabeled.

Given this possibility of confusion, words with final /p'/ are among the first to be misidentified and analyzed as if they did not have the palatalized gesture/feature. These are followed by lexical items with a final /t'/. The range of variation in the lexicon is quite limited, since all other items remain stable.

Let us now examine perception of derived items (stem + suffix: *apa*, *ap'a*, etc.). In these items all single consonants occur in intervocalic onset position. Thus, the decision about labeling these consonants as 'plain' or 'palatalized' is based on the same perceptibility scale as we used for the initial onset consonants in the previous case. Table 8.4 presents the surface forms in lexicons 1 to 4 as derived based on the perceptibility scale.

Table 8.4. Lexicons (surface forms) that can be derived based on the perceptibility scale for single consonants

Lexicon 0	Perceived as palatalized	Lexicon 1 *0.05*	Lexicon 2 *0.10*	Lexicon 3 *0.15*	Lexicon 4 *0.20*
at'a	0.96	at'a	at'a	at'a	at'a
ap'a	0.93	ap'a	ap'a	ap'a	ap'a
ata	0.12	ata	ata	ata	ata
apa	0.03	apa	apa	apa	apa

Note that since variation in perception of these items is minimal (from .03 to .12), they are correctly specified in terms of palatalization in all four lexicons.

Having labeled surface forms, the learner is ready to determine their underlying status, or generalize. In Lexicons 1 and 2 the original item *ap'* is variable, i.e. it can have surface representations [ap'] or [ap]. I assume that this decision may be made randomly (but see section 8.3.3). At the same time, the original item *ap'a* is always encoded as it is (i.e. [ap'a]). In one case (8.4a), the learner would have no evidence against positing an underlying representation as /ap'/, since the form of the morpheme is consistently [ap']. In the second case (8.4b), the learner, faced with two allomorphs, [ap] and [ap'-], has to decide between the two underlying forms, either /ap/ or /ap'/. Based on evidence from other items (e.g. [ap] and [apa], [at'] and [at'a], etc.), the learner is likely to arrive to the underlying form /ap'/ (8.4b).

(8.4)	Surface forms		Underlying form
a.	[ap']	[ap'-]	/ap'/
b.	[ap]	[ap'-]	/ap'/

Lexicons 3 and 4 show variability both in *ap'* ([ap']~[ap]) and *at'* ([at']~[at]). Applying the same procedure, the learner will posit the following correspondences between surface and underlying forms (8.5).

(8.5)

	Surface forms		Underlying form
a.	[ap']	[ap'-]	/ap'/
	[at']	[at'-]	/at'/
b.	[ap]	[ap'-]	/ap'/
	[at]	[at'-]	/at'/

Underlying forms for lexical items that do not show variation (e.g. *pa, p'a, ap, at*, etc.) are dealt with in the same manner as in (8.5a). Note that labeling the input *ap'* and *at'* as [ap] and [at] will result in homophony, since the other items *ap* and *at* are already present in the lexicon.

It is important, that in absence of alternations or other evidence an incorrectly labeled surface form can be classified by the learner as the true underlying form, thus leading to its re-analysis (Elan Dresher, p.c.).

Further, based on the lexical items (both their surface and underlying representations), the learner ranks, or constructs the grammar, a set of generalizations about how segments may be combined in the language. As shown in (8.6), the relative ranking of Ident[pal] against the same set of constraints on palatalized consonants gives three possible grammars. The arrow '»' stands for a crucial ranking of two constraints, while the absence of it shows that the constraints are unranked with respect to each other. The constraints on plain consonants can be assumed to be ranked below Ident[pal] and unranked with respect to each other or to the low ranked constraints on palatalized segments.

(8.6)

a. G1 **Ident[pal]** » *p'/_# *t'/_# *p'/#_ *t'/#_

b. G2 [redacted] » **Ident[pal]** » *t'/_# *p'/#_ *t'/#_

c. G3 [redacted] » **Ident[pal]** » *p'/#_ *t'/#_

The first grammar allows all of these segments in both positions, initial and final. The second one neutralizes labials in final coda, and the third one prohibits both palatalized labials and coronals in final environment. These grammars can derive the three attested distribution patterns discussed in Chapter 2. These patterns, Patterns 1, 2, and 3, are repeated in (8.7).

(8.7)

a. Pattern 1: Russian, Irish, Bulgarian (Nova Nadezhda)

	p	t	p'	t'
onset	*pa*	*ta*	*p'a*	*t'a*
final coda	*ap*	*at*	*ap'*	*at'*

b. Pattern 2: Russian (Cherdyn'), Belorussian, and Bulgarian (Krinichnoe)

	p	t	p'	t'
onset	*pa*	*ta*	*p'a*	*t'a*
final coda	*ap*	*at*		*at'*

c. Pattern 3: Russian (Vologda-Vyatka), Bulgarian (Standard)

	p	t	p'	t'
onset	*pa*	*ta*	*p'a*	*t'a*
final coda	*ap*	*at*		

The patterns attested in languages with a 3-way contrast (see section 2.1.3.1.1.1) are also easily generated, assuming that constraints on palatalized /p'/ are ranked above Ident[pal].

Thus a learner can arrive at three possible grammars and neutralization patterns without any *knowledge* of relative positional markedness or *awareness* of the difference between the two contexts or consonants in terms of their salience. This is accomplished simply by succeeding or failing at recovering gestures/features and constructing the grammars based on the posited lexical forms.

Recall from Chapter 2 that these patterns allowed us to make cross-linguistic generalizations and to posit implicational statements (repeated in (8.8)). These statements show asymmetries in terms position (onset vs. coda), palatalization (plain vs. palatalized), and place (labial vs. coronal). As we see now, these asymmetries are readily explainable by the derived perceptual scale based on a number of phonetic factors and its internalization by the grammar.

(8.8) a. Environment: onset vs. final coda

C vs. C'/_# > C vs. C'/_V
• If a language maintains a plain-palatalized contrast in the final coda, it also has it in onset;

b. Consonants: plain vs. palatalized

C'/_# > C/_#
• If a language has a palatalized consonant in the final coda, it also has plain consonants in this position;

c. Consonants: labial vs. coronal

C_{lab}/_# vs. C_{lab}/_# > C_{cor}/_# vs. C'_{cor}/_#
• If a language maintains a plain-palatalized contrast in labials in the final coda, it also has it in coronals in this position.

8.2.4 Constraints on coda consonants in clusters

In this section I discuss cases in which the input to the learner consists of words with coda consonants in clusters. In Language X these are derived by combination of VC and CV stems, where C stands for any of the four consonants (/p/, /t/, /p'/, /t'/) and V is /a/.

8.2.4.1 Phonetic factors

Recall that in the coda environment, articulatory differences between plain and palatalized consonants in terms of the Tongue Body gesture are relatively small and to a large degree affected by the quality (palatalization and place) of the following segment. Thus, the resulting VC transition is less informative. There is no direct information about the release of the first consonant due to the closure of the second consonant. At the same time, the overlap of two primary gestures (C_1 and C_2) often leaves C_1 without important burst information. This is most common in homorganic clusters and in some hetero-organic clusters, depending on place of articulation.

Again, the palatalized labial is at a greater disadvantage due to its articulatory properties (lower Tongue Body timed at the release of primary gesture) and acoustic (release burst quality) properties. Overall, the differences between /p/ and /p'/ in this position are much smaller than in the onset environment.

All this (with additional auditory factors) results in poor recoverability of the coda consonants. These are almost unrecoverable in noise. Figure 8.3 shows the scales for plain and palatalized labials (a) and coronals (b) where these consonants in coda (C_1) are perceived as palatalized, based on the combined values for three conditions (from Table 8.1). Notice the striking difference from the previously discussed relatively robust distinction between the two categories (plain and palatalized). Here the difference is much more gradient, and its overall range is smaller. Labials show more bias towards the plain category than coronals do. Place confusion (not shown here) due to the frequent absence of a release burst adds to the overall recoverability problem.

Of the two contexts, before plain and before palatalized segments, the latter affects the perception of C_1 the most.

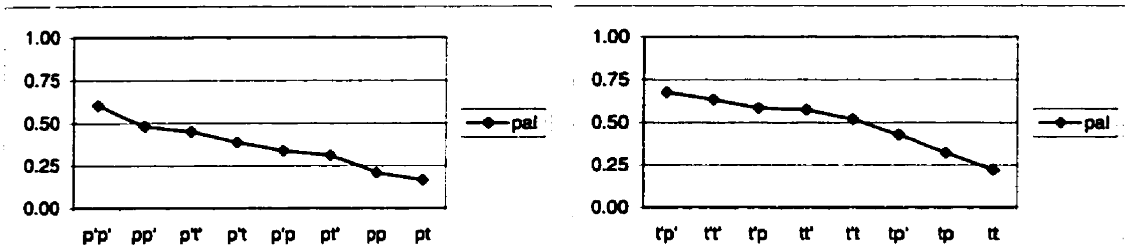


Figure 8.3 Perception of labials (a) and coronals (b) in terms of palatalization, coda consonants in clusters

As I show below, this strongly affects what surface and underlying forms are specified in the lexicon and what the resultant grammars are. I discuss the contexts before plain and palatalized consonants separately.

8.2.4.2 Possible lexicons and grammars: _C

In Table 8.5 the lexical items (VCCV and VC'CV) are presented in terms of perceived palatalization, from the highest score to the lowest. The horizontal line marks the division line between the two categories (.50).

Table 8.5 Lexicons (surface forms) that can be derived based on the perceptibility scale for coda consonants in clusters before a plain consonant

Lexicon 0	Perceived as palatalized	Lexicon 1 *0.05*	Lexicon 2 *0.10*	Lexicon 3 *0.15*	Lexicon 4 *0.20*
at'pa	0.58	at'pa			
at'ta	0.52				
ap'ta	**0.39**				
ap'pa	**0.34**				
atpa	0.32	atpa	atpa	atpa	
atta	0.22	atta	atta	atta	atta
appa	0.21	appa	appa	appa	appa
apta	0.16	apta	apta	apta	apta

Notice that all the items with palatalized C_1 are very close to the category border line, and often below it (given in bold), in the category of 'plain'. Thus we are to expect that our learner will show substantial variability with these items. The confusion is most likely to be in the words with palatalized labials, as these are well below the category boundary: *ap'pa* (a score of .34) and *ap'ta* (.39). Thus, out of 100 tokens of *ap'pa* only 34 are be perceived with a palatalized /p'/ in coda (i.e. [ap'pa]). The rest are identified with a plain /p/ ([appa]).

The next candidate for confusion is *at'ta* (.52), where /t'/ is before a homorganic consonant (Lexicons 1-4). It is followed by *at'pa* (.58). Notice that a tendency in the opposite direction is also possible; however, it is much less likely. The only item that involves variability is *atpa*, which is sometimes perceived and stored as [at'pa] (.32). The complete shift of /tp/ to /t'p/ is, however, less likely, since the cluster /t'p/ shows a higher degree of variability (Lexicons 2-4).

In contrast with the variable items, words with plain labials in coda (*apta*, *appa*) or /t/ before a homorganic plain segment (*atta*) are substantially more stable (.16-.22).

Overall, the perception of plain and palatalized contrasts before plain consonants induces a higher degree of variability in the lexicon compared to the previously discussed context.

Due to this degree of variation the task of generalizing across surface forms is more challenging. Let us consider the learner's strategy in identifying the underlying form of *ap'*- in the original items *ap'pa* and *ap'ta*. The chances of arriving at the correct underlying representation based on these two items only (8.9a) are not very high, since the items are often labeled incorrectly as [appa] and [apta]. Thus, the more likely scenario is to consider both [ap-] allomorphs as indicative of the underlying form /ap/ (8.9b).

(8.9)

	Surface forms		Underlying form
	+pa	*+ta*	
a.	[ap'-]	[ap'-]	/ap'/
b.	[ap-]	[ap-]	/ap/

It is mainly the reference to forms where /p'/ in this morpheme is a single consonant, *ap'a* and *ap'*, that would lead to the correctly posited form (8.10a). Recall, however, that the item *ap'* is also often labeled as [ap] (8.10b). The fact that most allomorphs do not show palatalization may eventually lead to the re-analysis of the underlying form to /ap/.

(8.10)

	Surface forms				Underlying form
	+pa	*+ta*	*+a*	*_#*	
a.	[ap-]	[ap-]	[ap'-]	[ap'-]	/ap'/
b.	[ap-]	[ap-]	[ap'-]	[ap-]	/ap'/

The same logic, although to a somewhat smaller extent, applies to the items with palatalized /t'/ in coda, *at'ta* and *at'pa*.

Based on Lexicons 1-4 (both their surface and underlying representations), the learner can posit a limited set of constraint rankings for this context. Those for palatalized consonants are shown in (8.11). I combine the two higher ranked environments (p'/_p and p'/_t) together, since the differences between them on the scale are minor.

(8.11)

a. G1	**Ident[pal]**	»	***p'/_p**		***p'/_t**		***t'/_t**	***t'/_p**
b. G2	[redacted]			»	**Ident[pal]**	»	***t'_t**	***t'/_p**
c. G3	[redacted]					»	**Ident[pal]** »	***t'/_p**
d. G4	[redacted]							**Ident[pal]**

All these rankings can generate the attested distribution patterns found cross-linguistically (Patterns 1-4). Notice that the scale also makes a prediction that, given this degree of confusion, the most 'faithful' to the input Grammar 1 (Pattern 1) is highly unstable, but not entirely impossible. A dialect of Bulgarian (Nova Nadezhda) is the only language in the sample that exhibits this pattern. Additional factors that may contribute to maintaining it are discussed in section 8.3.4. The neutralization of the contrast /t/ vs. /t'/ before a homorganic consonants (as in *at'ta*; Grammar 2, Pattern 2) is also rare, as predicted by the degree of variation in our lexicons. Most languages tend to maintain the more salient contrast (/t/ vs. /t'/) in the least affected context (_p) (Grammar 3, Pattern 3) or to neutralize the plain-palatalized contrast before plain consonants altogether (Grammar 4, Pattern 4).

Recall that Lexicon 4 showed some variability with respect to item *atpa* ([atpa] ~ [at'pa]). Interestingly, this is attested in Russian where neutralization of the distribution of /t/ and /t'/ before labials is almost complementary with respect to morphological environments: while /t'/ is found within stems, /t/ is common across prefix-stem boundaries (see section 2.2.2.2.1.3).

The markedness statements about the plain-palatalized contrast in this context (8.12) show asymmetries in terms environment (hetero-organic vs. homorganic), palatalization (plain vs.

palatalized), and place (labial vs. coronal). As in the previous case, these asymmetries follow from our derived perceptual scale and its interaction with the grammar.

(8.12)

a. Environment: hetero-organic ($_{ht}$) vs. homorganic ($_{hm}$)

C vs. C'/_C_{ht} > C vs. C'/_C_{hm}
• If a language maintains a plain-palatalized contrast (coronals) before a homorganic segment, it also has it in the hetero-organic context;

b. Consonants: plain vs. palatalized

C'/_C > C/_C
• If a language has a palatalized consonant before a plain segment, it also has a plain consonant in this environment;
c. Consonants: labial vs. coronal
p vs. p'/_C > t vs. t'/_C
• If a language maintains a plain-palatalized contrast before a plain segment in labials, it also has it in this position in coronals.

8.2.4.4 Possible lexicons and grammars: _C'

In this section I examine the context before palatalized consonants, the environment that is characterized by the lowest correct identification of segments.

In Table 8.6 the lexical items (VCC'V and VC'C'V) are presented in terms of perceived palatalization, from the highest score to the lowest. The values for two items, *att'a* and *ap't'a* are given in bold because they are either higher or lower than the border line between the two categories (.50).

Table 8.6 Lexicons (surface forms) that can be derived based on the perceptibility scale for coda consonants in clusters (before a palatalized consonant)

Lexicon 0	Perceived as palatalized	Lexicon 1 *0.05*	Lexicon 2 *0.10*	Lexicon 3 *0.15*	Lexicon 4 *0.20*
at'p'a	0.67	at'p'a	at'p'a	at'p'a	
at't'a	0.63	at't'a	at't'a		
ap'p'a	0.60	ap'p'a			
att'a	0.57				
app'a	0.48				
ap't'a	0.45				
atp'a	0.43	atp'a			
apt'a	0.31	apt'a	apt'a	apt'a	

Notice that, unlike before plain consonants, in this context the tendency to variability spreads in both directions. Plain consonants can be perceived as palatalized, or palatalized segments can be confused with plain. This correlates with homorganicity and place, with the most likely candidates for confusion *att'a* (with *at't'a*) and *ap't'a* (with *apt'a*), followed by *app'a*. Other lexical items are affected almost simultaneously in each direction. This leads to a situation where all the lexical items are variably represented. Thus, both words with either plain (e.g. *app'a*) or palatalized (*ap'p'a*) consonants in the same position show almost the same degree of variability.

The learner generalizing across surface forms has to inspect various contradictory pieces of evidence for one or the other underlying form. Let us consider the lexical items that have the morpheme *at-* as the first component of compounds *att'a* and *atp'a*. The first item is more often incorrectly labeled as [at't'a] rather than [att'a] (a score of .57); for the second one the correct labeling as [atp'a] is slightly more common over [at'p'a]. Thus, given three most likely sets of surface forms (8.13), the learner may either posit the correct underlying form /at/ (8.13ab), or may wrongly conclude that this form is /at'/ (8.13c). The latter solution can be avoided only if the learner refers to other surface forms with this morpheme (*at*, *atta*, and *atpa*).

(8.13)

	Surface forms		Underlying form
	+t'a	*+p'a*	
a.	[at-]	[at-]	/at/
b.	[at-]	[at'-]	/at/
c.	[at'-]	[at'-]	/at'/

Similar decisions have to be made about the other plain and palatalized consonants in this highly error-prone context.

I show the rankings of constraints that can theoretically result from these lexicons separately for plain (8.14) and palatalized (8.15) consonants as C_1. Recall that Ident[pal] refers to any of the two values of the feature, plain or palatalized.

(8.14)

a. G1 **Ident[pal]** » *t/_t' *p/_p' *t/_p' *p/_t'

b. G2 [redacted] » **Ident[pal]** » *p/_p' *t/_p' *p/_t'

c. G3 [redacted] » **Ident[pal]** » *t/_p' *p/_t'

d. G4 [redacted] » **Ident[pal]** » *p/_t'

e. G5 [redacted] » **Ident[pal]**

Grammar 1 corresponds to the situation found in Standard Bulgarian (Pattern 4 described in Chapter 2), where only plain consonants occur before the palatalized ones. Grammar 2 can result in the distribution of plain segments attested in Pattern 3 (Lithuanian dialects). Grammars 3 and 4 represent two varieties of Standard Russian (the more recent pronunciation and the prescribed norm). And finally, the last grammar (G5) neutralizes the plain segments in this context altogether, as in Pattern 2 (Irish and Lithuanian).

The constraints against the occurrence of plain segments are often counterbalanced by allowing palatalized segments instead (8.15). As a result the contrast between plain and palatalized segments in this environment is rarely maintained (only Russian: *atp'a* vs. *at'p'a*).

Now consider the ranking of constraints on palatalized segments. As in the previous cases, we see that the learner constructing a grammar is closely guided by what s/he has recovered, and, consequently, stored, labeled, and generalized. In this, situation, however, the choice of variable forms in the lexicon, and thus of relative rankings, is almost equally possible in both directions due to the perceptual bias both in favour of plain and palatalized segments and the relatively small difference between them.

(8.15)

a.	G1	**Ident[pal]**	»	***p'/_t'**		***p'/_p'**	***t'/_t'**	***t'/_p'**
b.	G2	[redacted]	»	**Ident[pal]**	»	***p'/_p'**	***t'/_t'**	***t'/_p'**
c.	G3	[redacted]				» **Ident[pal]**	» ***t'/_t'**	***t'/_p'**
d.	G4	[redacted]					» **Ident[pal]**	» ***t'/_p'**
e.	G5	[redacted]						» **Ident[pal]**

Thus, in rather commonly attested cases (Grammars 1 and 5) the learner appears to ignore minor differences and collapse all consonants and all environments together. For instance, i.e. *[+pal]/_[+pal] (Grammar 1) or *[-pal]/_[+pal] (Grammar 5). Recall that these languages (Irish, Lithuanian, and Bulgarian) apply the similarly general constraint before plain segments: *[+pal]/_[-pal] (i.e. generalizing it to *[αpal]/_[-αpal]). Other cases show more sensitivity to consonant (labial vs. coronal) and environment (before homorganic or hetero-organic consonant) differences.

In sum, we see that in cases where the phonetic differences and biases are minor and can be interpreted either way, languages show more variation and language-specific decisions about how constraints can be ranked.

Again, our model readily derives the markedness scales (8.16) attested for the plain-palatalized contrast before palatalized segments without positing these scales as an innate a part of UG.

(8.16) Environment: hetero-organic vs. homorganic (coronal)

a. C vs. $C'/\_C'_{hm}$ > C vs. $C'/\_C'_{ht}$
• If a language maintains a plain-palatalized contrast before a homorganic palatalized segment, it also has it in the hetero-organic context;

b. Consonants: labial vs. coronal

p vs. p'/_C' > t vs. t'/_C'
• If a language maintains a plain-palatalized contrast before a palatalized segment in labials, it also has it in this position in coronals.

c. Consonants: plain vs. palatalized
C'/_C > C/_C
• If a language has a plain consonant before a palatalized segment, it also has a palatalized consonant in this environment;

8.2.3 Constraints on onset consonants in clusters

Finally, I turn to the cases in which the learner identifies onset consonants in clusters (as part of stem 1 in VCCV compounds).

8.2.3.1 Phonetic factors

Recall that onset consonants, being maximally distinct from each other, show very little articulatory variability depending on quality of the consonant, or presence or absence of a preceding consonant. As a result listeners show almost 100% correct identification in terms of palatalization under all conditions. Figure 8.4 shows the scales for onset stops (C_2) in clusters where these consonants are perceived as palatalized, based on the combined values for three conditions (from Table 8.1). Labials (a) and coronals (b) are given separately.

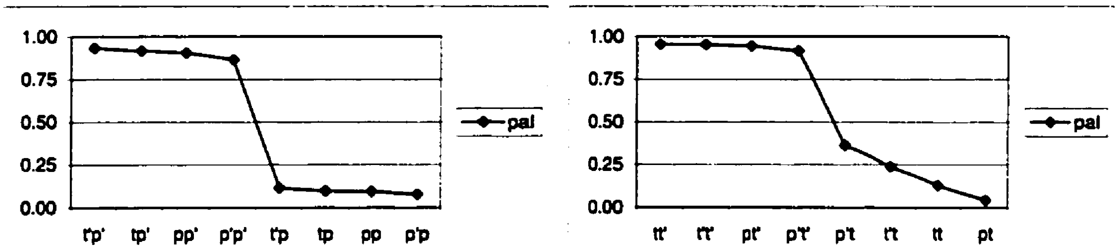


Figure 8.4 Perception of labials (a) and coronals (b) in terms of palatalization, onset consonants (C_2) in clusters

Notice that there is very little confusion in this environment, except for the possibility of misidentifying /t/ after palatalized consonants. Recall that the plain coronal is produced by fronting and some raising of Tongue Tip. The movement of this articulator, as a consequence, involves certain fronting of Tongue Body. In less clear cases (especially with noise) this can be perceived as an independent Tongue Body gesture/feature characteristic of a palatalized consonant. This is more likely in the proximity of an actual palatalized gesture, as in clusters /p't/ and /t't/.

8.2.3.2 Possible lexicons

The lexical items in Table 8.7 are ordered in terms of their perception as plain or palatalized. They are divided depending on the plain (a) or palatalized (b) quality of the preceding consonant.

Table 8.7 Lexicons (surface forms) that can be derived based on the perceptibility scale for onset consonants in clusters

a. After a plain consonant

Lexicon 0	Perceived as palatalized	Lexicon 1 *0.05*	Lexicon 2 *0.10*	Lexicon 3 *0.15*	Lexicon 4 *0.20*
att'a	0.95	att'a	att'a	att'a	att'a
apt'a	0.94	apt'a	apt'a	apt'a	apt'a
atp'a	0.92	atp'a	atp'a	atp'a	atp'a
app'a	0.91	app'a	app'a	app'a	app'a
atta	0.13	atta	atta	atta	atta
atpa	0.10	atpa	atpa	atpa	atpa
appa	0.09	appa	appa	appa	appa
apta	0.04	apta	apta	apta	apta

b. After a palatalized consonant

Lexicon 0	Perceived as palatalized	Lexicon 1 *0.05*	Lexicon 2 *0.10*	Lexicon 3 *0.15*	Lexicon 4 *0.20*
at't'a	0.95	at't'a	at't'a	at't'a	at't'a
at'p'a	0.93	at'p'a	at'p'a	at'p'a	at'p'a
ap't'a	0.92	ap't'a	ap't'a	ap't'a	ap't'a
ap'p'a	0.86	ap'p'a	ap'p'a	ap'p'a	ap'p'a
ap'ta	0.36	ap'ta	ap'ta	[redacted]	
at'ta	0.24	at'ta	at'ta	at'ta	at'ta
at'pa	0.12	at'pa	at'pa	at'pa	at'pa
ap'pa	0.08	ap'pa	ap'pa	ap'pa	ap'pa

Note that the only case that may cause a problem in the recovery of the plain and palatalized contrast in this environment is the item *ap'ta*. It is confused with *ap't'a* in Lexicons 3 and 4. The generalizing procedure for this case is shown in (8.17). Both the most likely set of surface forms (8.17a) and the less possible one (8.17b) are likely to lead the learner to positing the same correct underlying form /ta/. Additional evidence would come from the non-compound form *ta*.

(8.17) Surface forms — Underlying form

a. [ta], [-ta], [-ta], [-ta], [-ta] — /ta/

b. [ta], [-ta], [-ta], [-ta], [-t'a] — /ta/

Based on the lexicons the learner would always rank all of the constraints against onset consonants (except for *t/p'_ in some grammars) below the faithfulness constraint. Comparing these with each other and with those for single onsets can result in a more abstract constraint ranking as shown in (8.18). Here all the specific constraints on onset consonants are collapsed into two constraints on all palatalized ([+pal]) and plain ([-pal]) segments in the syllable onset position.

(8.18)

Ident[pal] » *[+pal]/Ons *[-pal]/Ons

This ranking is true for almost all of the languages that allow the contrast in both initial position and after a consonant in clusters, showing the same treatment for the two positions.

The exceptional case *t/p'_ » Ident[pal] is not attested in our sample. It should be noted, however, that a cluster of this type almost never occurs (except for the Nova Nadezhda dialect of Bulgarian) due to the restriction on /p'/ as C_1. A similar change of /k/ to /k'/ after palatalized consonants (mainly /t'/, /n'/, and /l'/) is widely attested in Russian dialects (Avanesov & Orlova 1965; Chapter 2). This process, often exhibited in synchronic alternations between /k/ and /k'/, is likely to be motivated by the same phonetic principles.

Note that unlike /t/, the plain labial is not affected by the preceding palatalized segments (Figure 8.4a). This correlates with relative stability of /t'p/ (*d'b, t'm, d'm*) clusters (West and East Slavic languages), where C_2 is labial.

The same patterning of onset consonants, whether single or in clusters, can be attributed to the fact that they are very similar perceptually (due to their articulatory and acoustic similarities).

8.2.4 Summary

As we saw the perceptual scales delimit the learner's choice of possible options each environment, resulting in a highly restricted set of synchronic lexicons and grammars. These grammars are able to generate all and only attested patterns and language types with respect to palatalization ((8.19) repeated from Chapter 2).

(8.19)

	_V *pa vs. p'a*	_# *ap vs. ap'*	_C *apta vs. ap'ta*	_C' *apt'a vs. ap't'a*
Type 1	yes	yes	yes	yes
Type 2	yes	yes	yes	
Type 3	yes	yes		
Type 4	yes			

Crucially, with reference to the process of gestural/featural recoverability and without the assumption of innateness of the markedness scales.

The model is also consistent with the attested course of sound change in palatalization as depalatalization or assimilatory palatalization in clusters (see Chapter 2). These types of change can be seen as a gradual shift over time from a less restrictive lexicon and corresponding grammar to more restrictive ones.

In our simplified simulation of acquisition we had to abstract away from a number of important factors. These are briefly addressed as issues for further research in the next section.

8.3 Outstanding issues

8.3.1 Other contexts and segments

We assumed in our simulation that all items of interest are pronounced in the same environment, between vowels. Obviously, this is far from the real running speech that a learner is exposed to. Each of the items can also occur either phrase-initially or finally, or before and after consonants. This can affect the recoverability of the items in either direction. Recall that phrase-final consonants are more likely to have a long and perceptually salient burst (especially /t'/) (see Chapter 3), and thus are in a better position to be correctly identified. At the same time, the segments before other consonants, especially the homorganic ones, are at a strong disadvantage in terms of recoverability of the contrast. We would expect recoverability, and thus the resulting grammars, to be affected by these considerations.

In our simplified model we assumed only one dimension (plain vs. palatalized) and did not consider the fact that consonants, especially final ones, are also very likely to be confused in terms of place, voice, and manner, as well as segmented in different ways (see Chapters 6 and 7). For instance, some of the tokens for the lexical item /ap'/ can be perceived and stored in a variety of ways: as [at'], [ak], [ajp], [ɛp], [af'], etc.). Based on our writing experiments, these errors are less likely than those in terms of palatalization alone. Nevertheless, this may also lead to additional confusion and biases. Recall that even when listeners correctly recovered the Tongue Body gesture/feature, they often had difficulties determining its relative extent and timing with respect to other gestures. This seems to be the motivation for the common changes /p'a/ > /pja/, /ap'/ > /ajp/ (see section 2.1.3). Attributing the Tongue Body raising to a front vowel is another common error that was a likely trigger of the change of the type /t'at'/ > /t'ɛt'/ or /t'ɛt/ (e.g. as in Russian dialects (Avanesov & Orlova 1965) or Upper Sorbian (Carlton 1990: 258; Chapter 2)).

In all of the experiments the vocalic context was limited to the low vowel /a/. It should be noted, however, that perception of palatalization is strongly affected by the nature of the preceding and following vowels (e.g. Kochetov 1999, Kavitskaya 2000). This factor and its consequences for phonotactic patterns of palatalization should be further investigated.

The details of articulation, acoustics, and perception of velar stops, as well as their phonotactics deserve a separate study. Consonants of other manners of articulation or voice specifications were beyond the scope of the current study. The languages of the current survey suggest that in general palatalized fricatives and nasals show patterning similar to the corresponding palatalized stops (although not without exceptions). Coronal approximants /r'/ and /l'/ tend to exhibit more deviation, being either less restricted (e.g. /l'/ in Russian and Belorussian) or more constrained (e.g. /r'/ in Ukrainian). Overall, palatalized labials tend to pattern as a class more often than palatalized coronals. This patterning in terms of natural classes is more consistent before plain consonants than before palatalized ones. All these questions are related to more general concept of relative similarity between segments and features and various sources for this similarity (cf. Frisch et al. 2000, Fleischhacker 2001). It appear that answers to these should also be sought in a combination articulatory, acoustic, and cognitive properties of those segments.

8.3.2 Language-particular phonetic differences

Since the focus of the thesis is primarily on Russian, articulatory and acoustic properties of plain and palatalized stops found in other languages were not experimentally investigated. It was assumed that these differences are not crucial for perception, the assumption at least in part confirmed by the fact that the Russian data allowed us to derive the universal patterns. This, however, does not imply that these differences (e.g. Tongue Body magnitude and duration, location of the primary coronal constriction, the degree of secondary velarization or labialization, etc.) are not important. A careful analysis of cross-linguistic articulatory, acoustic parameters involved in plain palatalized contrast as well as their perception is an important avenue for future research.

8.3.3 Frequency

While the role of lexical or phoneme frequency in a synchronic grammar may not as obvious (however, see Frisch et al. 2000, Pierrehumbert 2001), it is likely to be a major part of the acquiring the language phonotactics. In our acquisition 'experiment' the frequency of the lexical items, and thus of the segments, was the same (100 tokens of each item and 700 tokens of each segment). Clearly, this is not the case in any natural language. If language X had frequencies of segments similar to Russian, the occurrence of the final /p'/ would be 20 times lower than that of its plain counterpart, /p/ (based on the corpus of Russian, 8 tokens of /p'/ and 164 tokens of /p/; out of 33,000 words).

As demonstrated in Pierrehumbert 2001, the relative frequency of two phonemic categories (as shown in a simulation of a consonant lenition) has a direct impact on how new tokens are labeled. The fact that there are more stored instances of a more frequent category, and thus more cases to compare a new item against, leads to a systematic bias towards this category, particularly in less clear cases. The same mislabeling over time results incrementally in a drift of a less frequent segment category towards a more frequent one (both synchronically and historically; Elan Dresher, p.c.). This may remind us of the common depalatalization process in many languages with a plain-palatalized distinction (see Chapter 2).

Recall that the higher rate of neutralization before palatalized consonants in Russian is attested mostly frequent native vocabulary (see section 2.2.2.2.2.4). This is to be expected since more gestural overlap, no audible release, and thus poorer recoverability are often found in more frequent items that are more common in casual speech (Browman & Goldstein 1989; see also section 5.2.2.2).

Obviously, relative frequency of segments is to a large extent language-particular. Unlike in Russian, final /t'/ ([c]) in Czech is very rare (based on Kucera & Monroe 1968 it is 23 times less common than /t/; while in Russian final /t'/ is 1.7 times less common than final /t/ (Kochetov, in preparation)). We would expect the learner's bias towards /t/ (and thus the tendency towards neutralization of /t/ vs. /t'/) to be stronger in Czech than in Russian. This is indeed true, since many of the original final instances of /t'/ in Czech are depalatalized (Carlton 1990).

A careful consideration of frequency effects, should provide additional insights into the acquisition of phonotactics.

8.3.4 Lexicon, morphology, and alternations

In our model we assumed the traditional generative view of allomorphs as surface forms derived from a unique underlying form by application of phonological constraints. A number of issues raised in the language survey may not be readily accounted for by this view. Among these are a number of phenomena that fall into the class of 'paradigm uniformity' (e.g. Kurylowicz 1949, Steriade 1996).

Recall that presence of alternates where a palatalized segment is found before vowels (e.g. *ap'-a*) helps retain the contrast in other more marked contexts (e.g. *ap'* or *ap'ta*). It may even lead to the re-introduction of a category, as happened with final /t'/ and /d'/ in Czech (Carlton 1990). A lack of alternations can contribute to the loss of the final contrast (e.g. depalatalization of final nasal labials in non-alternating environments in Russian (Borkovskii & Kuznetsov 1965)).

The only attested case where all palatalized stops are allowed before plain consonants is Nova Nadezhda dialect of Bulgarian. In this language, however, the palatalized segments occur mainly (or possibly exclusively) in clusters that result from the addition of highly productive inflectional or derivational affixes (Khristov 1956).

Recall also that our analysis of the distribution of palatalized stops in Russian reveals that in most cases these occur across a stem-suffix boundary. It is rare that the contrast is maintained in monomorphemic words. This is in a sharp contrast with the situation at word boundaries, where usually no restrictions with respect to palatalization hold.

Models of lexicon that allow for all forms of lexical items to be represented (e.g. exemplar model of the lexicon (Johnson 1996, Bybee 2000, Pierrehumbert 2000)) may provide a more coherent explanation for these and other phenomena, hopefully without abandoning some of the important insights of the standard theory.

8.4 Summary and conclusion

Reference to a combination of phonetic articulatory, acoustic, and perceptual factors allows the model to make very specific predictions of what is a possible set of language grammars and what is the most and least likely environment for neutralization. This is done without any reference to prior, or innate, knowledge of positional markedness, or even prior knowledge of actual constraints. Instead, the learner's degrees of freedom are highly restricted by what s/he can recover from the signal.

In conclusion, phonetic and phonological universals and implicational hierarchies can largely be explained by reference to a range of physical, often extra-grammatical, factors (e.g. Ohala 1981, 1983, Kawasaki 1982, Lindblom, MacNeilage, & Studdert-Kennedy 1983, Maddieson 1984, among others). As we saw in this thesis, the cross-linguistic facts of the distribution of the plain-palatalized contrast and its neutralization follow the same pattern. In this view phonology is not entirely different from many other domains (chemistry, biology, sociology, etc.) where the principles of spontaneous emergence of order, or self-organization, have been identified as playing an important role (Kauffman 1995). The structures that are more stable (replicable and resistant to various kinds of pressures) are retained over time, while others are gradually discarded. In our case the difference between more or less stable phonetic/phonological structures is expressed in the gradient perceptual scales that we derived based on phonetic factors. These scales, however, are only raw input to the grammar. Articulatory and acoustic information about a contrast, recovered by a learner, is processed in a cognitive mode with its own principles and limitations. The grammar induces a certain arbitrariness between the phonetic input and its mental representation. As a result, the ranking of phonotactic constraints in the grammar will follow the general pattern of the perceptual hierarchies (derived from the corresponding factors), but will inevitably deviate from them in some arbitrary ways, generalizing across a number of linguistically relevant domains.

References

Abondolo, Daniel (ed.) 1998. *The Uralic languages.* New York, N.Y.: Routledge.

Akamatsu, Tsutomu. 1997. *Japanese phonetics: Theory and practice.* Muenchen: Lincom Europa.

Akamatsu, Tsutomu. 2000. *Japanese phonology: A functional approach.* Muenchen: Lincom Europa.

Ambrazas, Vytautas (ed.). 1997. *Lithuanian grammar.* Vilnius: Baltos Lankos.

Archangeli, Diana & Douglas Pulleyblank. 1994. *Grounded Phonology*, MIT Press, Cambridge.

Avanesov, R. I. 1972. *Russkoe literaturnoe proiznoshenie.* Moscow: Prosveshchenie.

Avanesov R.I. & S.V. Bromlei. 1986 (eds.) *Dialektologicheskii atlas russkogo iazyka (DARI).* Moscow: Akademia nauk SSSR.

Avanesov, R.I. & Orlova V.G. 1965. *Russkaia dialektologiia.* Akademiia nauk SSSR. Institut russkogo iazyka. Moskva: Nauka.

Avery, P. & K. Rice. 1989. Segment structure and coronal underspecification. *Phonology* 6(2): 179-200.

Azarkh, Y. S. 1967. In: L. L. Kasatkin et al. (eds.). *Ocherki po fonetike severnorusskikh govorov.* Moscow: Nauka.

Azarkh, Y. S. 1973. Otverdenie parnykh miagkikh soglasnykh pered glasnymi v vologodsko-kirovskikh govorakh. In: Bromlei, S.V. et al. (eds.), *Issledovania po russkoi dialektologii.* Moscow: Nauka, pp.

Ball, Martin J. & James Fife (eds.) 1993. *The Celtic languages.* New York, N.Y.: Routledge.

Basara, A., Basara, J., Wójtowicz, J., & H. Zdunska. 1959. *Studia fonetyczne z Warmii i Mazur. I. Konsonantyzm.* Wroclaw, Poland: Wydawnictwo Polskiej Akademii Nauk.

Bethin, Christina Y. 1992. *Polish syllables: the role of prosody in phonology and morphology.* Columbus, Ohio: Slavica.

Bhat, D. N. S. 1978. A General Study of Palatalization. In J. H. Greenberg (ed.), *Universals of Language*, Vol. 2: Phonology. Stanford, California: Stanford UP. 47-91.

Bladon, A. 1986. Phonetics for hearers. In G. McGregor (ed.), *Language for hearers.* Oxford: Pergamon, 1-24.

Bolla, K. 1981. *A conspectus of Russian speech sounds.* Cologne: Boelau.

Bondarko, L.V. 1981. *Foneticheskoe opisanie iazyka i fonologicheskoe opisanie rechi.* Moscow: Nauka.

Bondarko, L.V. 1977. *Zvukovoi stroi sovremennogo russkogo iazyka.* Moscow: Prosveshchenie.

Borkovskii V.I. & P.S. Kuznetsov. 1965. *Istoricheskaia grammatika russkogo iazyka.* Moscow: Nauka.

Bregman, A.S. 1990. *Auditory scene analysis: The perceptual organization of sound.* Cambridge, MA: MIT Press.

Browman, C. P., & Goldstein, L. 1986. Towards an articulatory phonology. *Phonology Yearbook*, 3, 219-252.

Browman, Catherine P. & Louis Goldstein 1989. Articulatory gestures as phonological units. *Phonology* 6. 201-252.

Browman, C. P., & Goldstein, L. 1990a. Gestural specification using dynamically-defined articulatory structures. *Journal of Phonetics*, 18, 299-320.

Browman, C. P., & Goldstein, L. 1990b. Tiers in articulatory phonology, with some implications for casual speech. In T. Kingston & M. E. Beckman (eds.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech.* Cambridge University Press. 341-376.

Browman, C. P., & Goldstein, L. 1992. Articulatory phonology: An overview. *Phonetica*, 493-4. 155-180.

Browman, C. P., & Goldstein, L. 1995a. Dynamics and articulatory phonology. In T. van Gelder & R. F. Port (eds.), *Mind as Motion*. Cambridge, MA: MIT Press. 175-193.

Browman, C. P., & Goldstein, L. 1995b. Gestural syllable position effects in American English. In F. Bell-Berti & L. J. Raphael (eds.), *Producing speech: Contemporary issues*. For Katherine Safford Harris. Woodbury, NY: AIP Press. 19-34.

Browman, Catherine P. & Louis Goldstein. 1999. Competing constraints on intergestural coordination and self-organization of phonological structures. In: *Bulletin de la Communication Parlée*. Grenoble.

Byrd, D. 1992 Perception of assimilation in consonant clusters: A gestural model. *Phonetica* 49, 1-24.

Carlton, Terence R. 1990. *Introduction to the phonological history of the Slavic languages*. Columbus, Ohio: Slavica.

Chekman, V.M. 1970. *Gistoryja protsipastaulenniau pa tsverdostsi-miakkastsi u belaruskai move*. Minsk: Navuka i tekhnika.

Chitoran I., Goldstein L., Byrd, D. (in press) Gestural overlap and recoverability: Articulatory evidence from Georgian. *Papers in laboratory phonology, 7.*

Chomsky, Noam. & Morris Halle. 1968. *The sound pattern of English*. New York: Harper Row.

Clements, G. N. 1985. The geometry of phonological features. *Phonology Yearbook* 2, 225-252.

Clements G. N. & E. Hume. 1995. The internal organization of segments. In: J. Goldsmith (ed.) *The handbook of phonological theory*. Blackwell Publishers, 245-306.

Cohen J.D., MacWhinney B., Flatt M., & Provost J. 1993. PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers*, 25(2), 257-271.

Comrie, Bernard & Greville G. Corbett, eds. 1993. *The Slavonic languages,*. London, New York, N.Y.: Routledge.

De Burca, Sean 1958. *The Irish of Tourkmakeady, Co. Mayo: A phonemic study*. Dublin: Dublin Institute for Advanced Studies.

Decsy, Gyula 1966. *Yurak chrestomathy*. Bloomignton, Ind.: Indiana University.

Dejna, Karol. 1993. *Dialekty polskie*. Wroclaw, Poland: Zaklad Narodowy Im. Ossolinskich Wydawnictwo.

Derkach, M. 1975. Acoustic cues of softness in Russian syllables and their application in automatic speech recognition. In: *Auditory analysis and perception of speech*, G. Fant & M.A.A. Tatham (eds.). New York: Academic Press. 349-358.

Dmitrienko, S.N. 1985. *Fonemy russkogo iazyka*. Moscow: Nauka.

Dresher, Elan. 1999. Charting the learning path: Cues to parameter setting," *Linguistic Inquiry* 30, 27-67.

Erofeeva, E.V. 1997. *Eksperimental'noe issledovaniefonetiki regional'nogo varianta literaturnogo iazuka*. Perm': Perm' University.

Fleischhacker, Heidi. 2001. Experimental results on relative similarity. Paper presented at 2001 annual meeting of Linguistic Society of America. January 2001.

Flemming, Edward Stanton. 1995. Auditory representations in phonology. Doctoral dissertation. University of California, Los Angeles.

Fowler , Carol A. 1986. An event approach to the study of speech perception from a direct—realist perspective. *Journal of Phonetics*, 14, 3-28.

Fowler, Carol A. 1996. Listeners do hear sounds, not tongues. *Journal of Acoustic Society of America*, 99 (3)..

Frisch, S.A., Large, N.R., & Pisoni, D.B. 2000. Perception of wordlikeliness: Effects of segment probability and length on the processing of nonwords. *Journal of memory and language* 42. 481-496.

Gafos, A. I. 2000 The grammatical basis of gestural coordination. Ms., New York University.

Gick, Brian. 1999. The articulatory basis of syllable structure: A study of English glides and liquids. Doctoral dissertation. Yale University, New Haven, CT.

Hale, Mark & Charles Reiss. 2000. Substance abuse and dysfunctionalism: Current trends in phonology. *Linguistic Inquiry 31:1.*

Halle, Morris. 1959. *The sound pattern of Russian.* The Hague: Mouton.

Hamilton, William S. 1980. *Introduction to Russian phonology and word structure.* Columbus, Ohio: Slavica.

Hamilton, Philip James. 1996. Phonetic constraints and markedness in the phonotactics of Australian aboriginal languages. Doctoral thesis. University of Toronto.

Hayes, Bruce. 1999. Phonological acquisition in Optimality Theory: The early stages. Ms. UCLA.

Hume, Elizabeth. 1992. *Front vowels, coronal consonants and their interaction in non-linear phonology.* Published Ph.D. dissertation. Ithaca, N.Y.: Cornell University.

Hume, Elizabeth. 2001 (to appear). Metathesis: Formal and functional considerations. In: Elizabeth Hume, Norval Smith & Jeroen van de Weijer (eds.). Surface syllable structure and segment sequencing. HIL occasional papers. Leiden NL: HIL.

Hume, Elizabeth, Keith Johnson, Misun Seo, & Georgios Tserdanelis. 1999. A cross-linguistic study of stop place perception. In: *Proceedings of the 14th International Congress of Phonetic Sciences* (ICPhS99). 2069-2072.

Hume, Elizabeth & Keith Johnson. 2001. A model of the interplay of speech perception and phonology. In Elizabeth Hume & Keith Johnson (eds.), The role of speech perception in phonology. New York: Academic Press. 3-26.

Iartseva, V. N. 1993. *Iazyki mira. Ural'skie iazyki.* Moscow: Nauka.

Jakobson, Roman. 1971. K kharakteristike evraziiskogo yazykovogo soyuza. *Roman Jakobson: Selected Writings.* vol. 1: Phonological Studies. The Hague: Mouton. 144-201.

Jakobson, R, G. Fant, & M. Halle. 1951. *Preliminaries to speech analysis.* Acoustics Laboratory, MIT, Cambridge, MA.

Johnson, Keith. 1996. Speech perception without speaker normalization. In Johnson, K. & Mullenix (eds.) *Talker variability in speech processing.* San Diego: Academic Press.

Johnson, Keith. 1997. *Acoustic and auditory phonetics.* Cambridge, Mass.: Blackwell.

Jones, Daniel & Dennis Ward. 1969. *The phonetics of Russian.* Cambridge: University Press.

Jun, J. 1995. Perceptual and articulatory factors in place assimilation: An Optimality Theoretic approach. PhD Dissertation. UCLA.

Kauffman, S. 1995. *At home in the universe: The search for the laws of self-organization and complexity.* Oxford: Oxford University Press.

Kavitskaya, Darya. 2000. Perceptual salience and palatalization in Russian. Ms. University of California at Berkeley.

Kawasaki, Haruko. 1982. An acoustic basis for universal constraints on sound sequences. Doctoral dissertation. University of California, Berkeley.

Kelso, J. A. S., Saltzman, E. L., & Tuller, B. 1986. The dynamical perspective on speech production: Data and theory. *Journal of Phonetics, 14*(1). 29-59.

Khristov, Georgi. 1956. Izvestiia na instituta za bulgarski jezik. vol. 4. Govorut na s. Nova Nadezhda, Khaskovsko. Sofia: Izdanie na bulgarskata akademiia naukite. 177-253.

Kent, D. Ray and Charles Read. 1992. *The acoustic analysis of speech.* San Diego, CA: Singular Publishing Group.

Kingston, J. & R. Diehl. 1994. Phonetic knowledge. *Language* 70, 419-454.

Kiparsky, Valentin. 1979. *Russian historical grammar.* Vol.1. The development of the sound system. Ann Arbor, MI: Ardis.

Kochetov, Alexei. 1999a. Testing Licensing by Cue: A case of plain-palatalized coronal contrast in Russian. Ms., University of Toronto.

Kochetov, Alexei. 1999b. Social change and language change in a rural Russian community. Toronto Working Papers in Linguistics.

Kochetov Alexei & Louis Goldstein. 2001. Competing recoverability factors and inter-gestural phasing in Russian stop clusters. Paper presented at 2001 annual meeting of Linguistic Society of America. January 2001 (slides available at: http://www.chass.utoronto.ca/akocheto~).

Kucera, Henry & George K. Monroe. 1968. *A comparative quantitative phonology of Russian, Czech, and German.* New York: American Elsevier.

Kuznetsova, A.M. 1965. *Izmeneniia glassnykh pod vliianiem sosednikh soglasnykh.* Moscow: Nauka.

Kuznetsova, A.M. 1969. Nekotorye voprosy foneticheskoi kharakteristiki iavlenia tverdosti-miagkosti soglasnykh v russkikh govorakh. In: *Eksperimental'no-foneticheskoe izuchenie russkikh govorov*, S.S. Vysotskii (ed.). Moscow: Nauka. 35-140.

Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages.* Cambridge, Massachusetts: Blackwell.

Leskinen, H. 1984. Karelischen. In *Studien zur phonologischen Beschreibung uralischen Sprachen*, P. Hajdu & Laszlo Honti (eds.). Budapest: Akadémiai Kiad6. 247-257.

Liberman, A. M. 1996. *Speech: A special code.* Cambridge, Mass.: The MIT Press.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. 1967. Perception of the speech code. *Psychological Review, 74*, 431-461.

Liberman, A. M., & Mattingly, I. G. 1985. The motor theory of speech perception revised. *Cognition, 21*. 1-36.

Lindblom, B.E., P. MacNeilage, & M. Studdert-Kennedy. 1983. Self-organizing processes and the explanation of language universal. In: B. Butterworth, B. Comrie, & O. Dahl (eds.), *Explanations for language universals.* The Hague: Mouton. 181-203.

Lytkin, V. I., (ed.) 1962. *Komi-Permiatskyi iazyk.* Kudymkar: Komi-Permiatskoe knizhnoe izdatel'stvo.

MacAulay, Donald (ed.) 1992. *The Celtic languages.* New York, N.Y.: Cambridge University Press.

Maddieson, Ian: 1984. *Patterns of sounds.* Cambridge: Cambridge University Press.

Maddieson, Ian. 1997. Phonetic universals. In: *The handbook of phonetic sciences*, eds. William J. Hardcastle & John Laver (eds.). Cambridge, Mass.: Blackwell. 619-639.

Mayo, Peter. 1993. Belorussian. In: *The Slavonic languages:.* Bernard Comrie & Greville G. Corbett (eds.). London, New York : Routledge. 887-946.

Matusevich, M. I. 1976. *Sovremennyi russkii iazyk: fonetika.* Moscow: Prosveshchenie.

Ohala, J.J. 1981. The listener as a source of sound change. In: C.S. Masek, R.A. Hendrick, & M.F. Miller (eds.), *Papers from Chicago Linguistic Society parasession on language and behaviour.* Chicago: CLS. 178-203.

Ohala, J.J. 1983. The origins of sound patterns in vocal tract constraints. In: P.F. MacNeilage (ed.), *The production of speech.* New York: Springer Verlag. 189-216.

Ohala, J.J. 1990. The phonetics and phonology of aspects of assimilation. In J. Kingston & M.E. Beckman (eds.), *Papers in laboratory phonology I: Between the grammar and physics of speech.* Cambridge: Cambridge University Press.

Orlova, Varvara Georgievna (ed.). *Obrazovanie severnorusskogo narechiia i srednerusskikh govorov.* Moskva, Nauka, 1970.

Pierrehumbert, J. 2001. Exemplar dynamics: Word frequency, lenition, and contrast. In Bybee, J. & Hopper, P. (eds.). *Frequency effects and emergent grammar.* Amsterdam: John Benjamins.

Pierrehumbert, J., Beckman, M. E & Ladd, D. R. (in press) Conceptual foundations of phonology as a laboratory science. In Burton-Roberts, N., Carr, P. & Docherty, G. (eds) *Phonological knowledge*, Oxford, UK: Oxford University Press.

Prince, Alan & Paul Smolensky 1993. *Optimality Theory: Constraint interaction in generative grammar*. Manuscript, Rutgers University.

Purcell, Edward T. 1979. Formant frequency patterns in Russian VCV utterances. *Journal of Acoustic Society of America*, 66(6). 1691-1702.

Redei, K. 1984. Erza Mordwinisch. In *Studien zur phonologischen Beschreibung uralischen Sprachen*, Hajdu, P. & Laszlo Honti (eds.). Budapest: Akadémiai Kiadó. 247-257.

Redford, Melissa A. & Randy Diehl. 1999. The relative perceptual distinctiveness of initial and final consonants in CVC syllables. *Journal of Acoustic Society of America*, 106(3). 1555-1565.

Scatton, Ernest. 1993. Bulgarian. In: *The Slavonic languages:*. Bernard Comrie & Greville G. Corbett (eds.). London, New York : Routledge. 188-248.

Shevelov, George. 1993. Ukrainian. In: *The Slavonic languages:*. Bernard Comrie & Greville G. Corbett (eds.). London, New York : Routledge.

Short, David. 1993a. Czech. In: *The Slavonic languages:*. Bernard Comrie & Greville G. Corbett (eds.). London, New York : Routledge. 455-532.

Short, David. 1993b. Slovak. In: *The Slavonic languages:*. Bernard Comrie & Greville G. Corbett (eds.). London, New York : Routledge. 533-592.

Shtern, A.S. & E.V. Erofeeva. 1998. Fonetika govorov i literaturnogo iazyka na territorii permskoi oblasti. In: Erofeeva E.V. (ed.) *Russkie govory permskogo regiona: Formirovanie i razvitie*. Perm: Izdatel'stvo permskogo universiteta. 31-133.

Silverman, Daniel. 1997. *Phasing and recoverability*. New York: Garland.

Skalozub, L.G. 1963. *Palatogrammy i rentgenogrammy soglasnykh fonem russkogo literaturnogo iazyka*. Kiev: Izdatel'stvo Kievskogo universiteta.

Steriade, Donca. 1997. Phonetics in phonology: The case of laryngeal neutralization. Ms., University of California, Los Angeles.

Steriade, Donca. 2001. Directional asymmetries in place assimilation: a perceptual account. In: Elizabeth Hume & Keith Johnson (eds.), *The role of speech perception in phonology*. New York: Academic Press. 219-250.

Stieber, Zdzislaw. 1973. *A historical phonology of the Polish language*. Heidelberg: Winter.

Stone, Gerald. 1993a. Sorbian. In: *The Slavonic languages:*. Bernard Comrie & Greville G. Corbett (eds.). London, New York : Routledge. 593-685.

Stone, Gerald. 1993b. Cassubian. In: *The Slavonic languages:*. Bernard Comrie & Greville G. Corbett (eds.). London, New York : Routledge. 759-794.

Studdert-Kennedy, M. 1998. The particulate origins of language generativity. In Hurford, J. Studdert-Kennedy, M. & Knight, C. (eds). *Approaches to the evolution of language*. Cambridge: Cambridge University Press: 202-221.

Surprenant, A. & Goldstein, L. 1998. The perception of speech gestures. *Journal of the Acoustical Society of America*. 104, 518-529.

Tiede, Mark K., Eric Vatikiotis-Bateson, Phil Hoole, Hani Yehia. 1999. Magnetometer data acquisition and analysis software for speech production research. *ATR Technical Report TR-H*. ATR Human Information Processing Labs, Japan.

Timberlake, Alan: 1993. Russian. In: *The Slavonic languages:*. Bernard Comrie & Greville G. Corbett (eds.). London, New York : Routledge. 827-886.

Townsend, Charles E. 1980. *Russian word-formation*. Columbus: Slavica.

Trubetzkoy, N. S. 1958/1969. *Principles of phonology*. Berkeley, CA: University of California Press.

Vance, Timothy J. 1987. *An introduction to Japanese phonology*. New York: SUNY Press.

Wenck, Guenther. 1966. The phonemics of Japanese: questions and attempts. Wiesbaden O. Harrassowitz.

Zalizniak, A.A. 1977. *Grammaticheskii slovar' russkogo iazyka: Slovoizmenenie*. Moscow: Russkii Iazyk.

Zdunska, Helena. 1965. *Studia nad fonetyka gwar mazowieckich: Konsonantyzm.* Wroclaw: Wydawnictwo Polskiej Akademii Nauk.

Zsiga, E. C. 1995. An acoustic and electropalatographic study of lexical and postlexical palatalization in American English. In B. Connell & A. Arvaniti (eds.), *Phonology and phonetic evidence: Papers in laboratory phonology IV*. Cambridge: Cambridge University Press. 282-302.

Zsiga E.C. 2000 Phonetic alignment constraints: Consonant overlap and palatalization in English and Russian. *Journal of Phonetics*, 28 (1). 69-102.