Palatalization and Utilization of Contrast: An Information-theoretic Investigation of Palatalization in Russian

Thesis

Presented in Partial Fulfillment of the Requirements for the Degree Master of Arts in the Graduate School of The Ohio State University

By

Jeff Parker, B.A., M.A.

Graduate Program in Linguistics

The Ohio State University

2013

Thesis Committee:

Brian Joseph, Co-Advisor

Andrea D. Sims, Co-Advisor

Cynthia Clopper
Abstract

Traditionally phonological analysis separates sounds into two categories: those that are contrastive, i.e. can distinguish words, and those that are not, i.e. are predictably distributed or vary without changing meaning. This binary definition of contrast, however, falls short in dealing with a number of phonological relationships in which sounds exhibit some traits of contrastive sounds, as well as traits associated with non-contrastive sounds. Such intermediate phonological relationships vary in the extent to which they utilize contrast along two dimensions: contexts in which the contrast is utilized and number of lexical items that utilize the contrast. I argue that phonology must include not only the potential, but the utilization of the potential to distinguish words in order to fulfill the implicit expectations built around the concept of contrast. Investigating the utilization of contrast as a quantifiable aspect of phonological relationships allows the phonologist to classify phonological relationships more accurately and more directly predict their effects on the phonological system. In this thesis I use Hall's (2009) Probabilistic Phonological Relationship Model, which employs the information theoretic concepts of probability and entropy, to quantify the degree to which sounds utilize contrast in specific environments. I look at the contrast between palatalized and non-palatalized consonants represented in three sets of Russian data: velars before non-front vowels, a set of consonants word-
finally and a set of consonants before palatalized consonants. Each set of data supports
the need to investigate utilization of contrast as a core trait of phonological relationships.

I show that velars before non-front vowels minimally utilize contrast in two ways:
they exhibit contrast in only one context, and only a limited number of lexical items
exhibit the contrast in that context. I also show that contrastive consonants for which both
palatalized and non-palatalized consonants exist with the same voicing, place
and manner of articulation (paired consonants), vary greatly in the extent to which they
utilize their contrasts word-finally. I show that this gradience has clear parallels in
perception and production, suggesting that utilization of contrast, not just contrast itself,
is a key factor for speakers. Labials, which exhibit a low degree of contrast word-finally,
suffer from low levels of accuracy during perception and where a palatalized labial is
expected a non-palatalized labial is sometimes produced. Lastly, I show that the level of
contrast between palatalized and non-palatalized dentals before palatalized dentals and
labials has been changing in Russian. All dentals are said to have been palatalized before
palatalized dentals and labials, an effect of regressive palatalization. However, regressive
palatalization has declined in contemporary Russian, resulting in a new possibility of
contrast for dentals before palatalized dentals and labials. I present new empirical data
which suggests that the majority of targets of regressive palatalization are no longer
palatalized in productions of contemporary speakers. This suggests that while the
previous neutralization of contrast in this context has been eroded, the contrast has not
been utilized systematically, as it is in other contexts, e.g. word-finally. Instead, the
contrast between palatalized and non-palatalized dentals before palatalized dentals and
labials have shifted from being neutralized to being (largely) non-utilized in this context.
Together these three sets of data suggest that the utilization of contrast is a key component of defining phonological relationships and their effects on speakers. Without an explicit notion of utilization of contrast the synchronic and diachronic effects of contrast discussed in this thesis cannot be accounted for.
Acknowledgments

I owe much of my growth as an individual and development as a scholar to friends, colleagues and family members. I am particularly indebted to Andrea Sims for countless hours of conversation during which we discussed all aspects of this project. The things I have learned in working with Andrea stretch far beyond the topics discussed in this thesis. I have learned that her theoretical comments as well as professional advice are invaluable and I thank her for giving such advice kindly, directly and generously.

I am also grateful to Brian Joseph and Cynthia Clopper for their help and comments. During the research for this thesis, I learned that the sooner I asked for and accepted the wise advice of my committee members, the sooner the ideas began to flow and the sooner the project moved in useful directions. I am also thankful for the comments of people who attended presentations of this work, including members of the Slavic Linguistics Forum at Ohio State University and attendees of the Seventh Annual Meeting of the Slavic Linguistics Society.

I am grateful to my wife, Amy, whom I have abandoned on a daily basis so that I could write this thesis, leaving her to take care of our family. She has been supportive in every aspect of my work and I would never have made it without her. I am also grateful for my kids, Lily, Kaleigh and Benson, who have each provided much needed distractions at various times. Taking a few minutes here and there to play a childish game
has often been the most productive time I spent writing. Every time I returned to my work with fresh eyes and uplifted morale.
Vita

2008 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . B.A., Linguistics and Russian, cum laude, Brigham Young University

2008-2009 . . . . . . . . . . . . . . . . . . . . . . . . . University Fellow, The Ohio State University

2010 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . M.A., Russian Linguistics, The Ohio State University

2011-2012 . . . . . . . . . . . . . . . . . . . . . . . . . Foreign Language and Area Studies Fellow, The Ohio State University

2009-2011; 2012-2013 . . . . . . . . . . . . . . . . . . Graduate Teaching Associate, The Ohio State University

Publications


Fields of Study

Major Field: Linguistics
# Table of Contents

Abstract .......................................................................................................................... ii  
Acknowledgements ....................................................................................................... v  
Vita ............................................................................................................................... vii  
Table of Contents ......................................................................................................... viii  
List of Tables ................................................................................................................ ix  
List of Figures ............................................................................................................. x  

Chapter 1: Introduction ............................................................................................... 1  
Chapter 2: Contrast in phonological theory ............................................................... 6  
Chapter 3: Quantifying utilization of contrast ............................................................ 19  
Chapter 4: ‘Paired’ consonants word-finally ............................................................. 29  
Chapter 5: Consonants before palatalized consonants; regressive palatalization ...... 36  
Chapter 6: Conclusion ............................................................................................... 55  

References .................................................................................................................... 58  

Appendix A: Overview of vowel phones and their co-occurrence with stress and (non-)palatalized consonants ................................................................................. 61
List of Tables

Table 1. Distribution of Russian velars before non-front vowels ................................. 22
Table 2. Probability for Russian velars in two contexts ............................................. 23
Table 3. Entropy for Russian velars in two contexts .................................................. 25
Table 4. An inventory of Russian consonants grouped according to palatalization ...... 30
Table 5. Entropy (based on word types) for paired consonants word-finally .......... 32
Table 6. Consonant clusters investigated ................................................................. 42
Table 7. Words with a high change in F2 ................................................................... 47
Table 8. Change in entropy as percent of word types that regressively palatalize changes for five consonants ................................................................. 50
List of Figures

Figure 1. Two dimensions of contrast utilization ................................................... 12
Figure 2. Decline of regression palatalization over time ........................................ 39
Figure 3. Comparison of change in F2 in control and targets clusters .................... 45
Figure 4. Change in F2 by consonant ................................................................. 48
Figure 5. Illustration of change in entropy during the decline of regressive palatalization ................................................................. 53
Chapter 1: Introduction

Russian consonants exhibit contrast for palatalization, a secondary place of articulation\(^1\). The robust nature of the contrast is evident in a large number of words before non-front vowels and word-finally, e.g. [m]at 'foul language' vs. [m\(^i\)]at 'crumpled', ma[t] 'foul language' vs. ma[t\(^i\)] 'mother'\(^2\). However, contrast between palatalized and non-palatalized consonants is heavily restricted in some contexts, e.g. before consonants. When a lateral precedes other consonants contrast is possible, evidenced by minimal pairs such as po[l\(^k\)]a 'polka (dance)' vs. po[lk]a 'shelf', though there are few such pairs. The contrast is limited for other consonants, if it exists at all. Panov (2002:113) asserts that (non-lateral) dentals were always palatalized before palatalized dentals and palatalized labials in the 19th century, e.g. če[t\(^v\)]ert 'one fourth', zapo[d\(^n\)]ik 'westerner'\(^3\). This was the result of the process of regressive palatalization in which consonants preceding palatalized consonants were palatalized (whether they would have been palatalized otherwise or not). Thus, regressive palatalization neutralized the contrast between palatalized and non-palatalized dentals before palatalized dentals and labials.

---

\(^1\) The majority of consonants in Russian exhibit this contrast; however, some consonants do not. For a full description of palatalization and Russian consonants see §4.1 below.

\(^2\) The contrast between palatalized and non-palatalized consonants is largely neutralized before front vowels. With few exceptions, only palatalized consonants occur before front vowels.

\(^3\) This claim is particular to "speakers of the literary language", particularly those in Moscow. Panov also discusses minor deviations from this norm that occur in other dialects/cities depending on the phonological context and the types of consonants in question. Thus, complete neutralization may have only applied to a relatively small set of speakers. In dialects/areas where there was not complete neutralization, it is still unlikely that consonants in this context could distinguish words.
However, the regularity of the process weakened in the following century, altering the possible combinations of palatalized and non-palatalized consonants. What had been impossible earlier, a non-palatalized dental occurring before a palatalized dental or labial, was becoming possible. According to Panov, this decline of regressive palatalization shifted the position of dentals before palatalized dentals and labials from weak (no contrast possible; neutralization) to strong (contrast possible) (126).

Additionally, Timberlake (2004:61) suggests that regressive palatalization has declined so much so that in contemporary Russian speakers have "rather less - if any - assimilation than was reported [previously]".

Together, these facts bring up a number of questions about this change and its effects on the phonological system. Panov's reclassification of dentals before palatalized dentals and labials suggests that contrast is possible in this context but gives no direct indication of whether the contrast is actually utilized in this context. As regressive palatalization declines in the language, to what extent does the contrast actually get utilized? As the process continues to decline, does the utilization of the contrast increase consistently and, if not, what other factors affect the direction of the change?

While each of these questions is important to understanding the system of contrasts in Russian, certain assumptions common to phonological theory limit the phonologist in addressing these questions. Most notably, traditional phonological analyses consider two sounds to be contrastive in a language if, in at least one context, they can distinguish lexical items. This definition presupposes that contrast is a binary phenomenon; it is either possible or impossible for two sounds to contrast in a given context. Panov's claims about dentals before palatalized dentals and labials clearly reflect
this definition of contrast: in this context dentals shifted from one category (no contrast) to the other (contrast possible). However, this reclassification of the contrast for dentals has unsatisfying implications. Classifying dentals in this context as simply 'contrastive' for palatalization implies that the contrast is systematically the same as the contrast for palatalization in other contexts, e.g. word-finally. It ignores the fact that few, if any, words utilize the contrast for palatalization of dentals before palatalized consonants while thousands of examples utilize the "same" contrast word-finally. Just because contrast is possible does not entail that the contrast is utilized.

Additionally, Panov's classification of dentals as contrastive for palatalization before palatalized dentals and labials provides no description of or prediction for the future of the contrast (and its effects on speakers) as regressive palatalization declines in the language. Consider two scenarios. In scenario 1 the number of words that utilize the contrast between palatalized and non-palatalized dentals before palatalized consonants increases until it equals the robust level to which dentals utilize the contrast elsewhere, e.g. word-finally. In this scenario, the classification of the context as contrastive for palatalized and non-palatalized dentals seems appropriate: the contrast between palatalized and non-palatalized dentals is robustly utilized to distinguish words, whether the dental occurs word-finally or before palatalized dentals and labials. In scenario 2, the process that neutralized contrast in this context earlier disappears, but the contrast never becomes utilized; only non-palatalized dentals occur. This scenario presents two problems for the traditional definition of contrast. First, it is unclear whether the potential to distinguish words exists because the contrast is not utilized, despite the fact that there is no motivating reason, e.g. neutralization, that prevents it from being utilized. Second, if
even one minimal pair utilizes the contrast, the context becomes classified as 'contrastive'
despite the obvious differences between it and other contexts where a similar contrast is
utilized more fully. In most respects the dentals in Scenario 2 are more similar to the
dentals whose contrast was neutralized than to dentals in contexts where the contrast is
more robustly utilized. In this scenario, there is not motivated reason to suggest that
palatalized dentals could not distinguish words; they merely do not.

Crucially, scenarios 1 and 2 are substantially different; however, if only the
potential to distinguish words is used in classifying the phonological relationships in each
scenario, there is no way to tease the two apart, nor is there a way to determine which
scenario more closely represents Russian. A traditional account does not measure the
extent to which a contrast is utilized in each context, making it incapable of identifying
whether the contrast becomes robustly utilized or becomes possible albeit not utilized,
making it more similar to neutralization than contrast for the speaker.

In this thesis I investigate contrast between palatalized and non-palatalized
consonants in Russian for three sets of data, corresponding to three different contexts:
velars before non-front vowels, paired consonants word-finally and consonants before
palatalized consonants. Each of these investigations provides a particular challenge for an
account that investigates only the potential for distinguishing words rather than the
utilization of such potential. As an alternative to investigating contrast in this binary way,
I follow Hall's Probabilistic Phonological Relationship Model (2009), which employs
information theoretic tools to measure the utilization of a contrast in a (context-dependent
and) gradient way. I argue that investigating the utilization of contrast in this way gives
the phonologist much needed tools to describe intermediary phonological relationships and historical changes, as well as their effects on the phonological system.

In Chapter 2 I discuss contrast in phonology, including two dimensions along which intermediate relationships exist and why these relationships are problematic for defining contrast only as the potential to distinguish words. In Chapter 3 I explain how the utilization of contrast can be measured in gradient terms using Hall's (2009) Probabilistic Phonological Relationship Model and illustrate the approach using Russian velars. I then present two additional sets of data that illustrate the need for investigations of contrast based on the gradient extent to which it is utilized: a set of consonants word-finally (Chapter 4) and a set of consonants before palatalized consonants (Chapter 5). I end with my conclusions (Chapter 6).
Chapter 2: Contrast in phonological theory

2 Contrast and the phoneme

The insight that sounds in a language can be mapped from an infinitely variable acoustic signal to a set of discrete categories (phonemes) is crucially predicated on the idea of contrast. Sounds which can distinguish words are unpredictably distributed and therefore must be specified at some abstract level, e.g. as phonemes, the discovery of which is "one of the most magnificent achievements of linguistic science" (Krámský 1974:7) and "the great organizing principle of 20th century phonology" (Goldsmith 1998:7).

The existence of phonemes is predicated on contrast that exists on a level distinct from surface phonetics. For example, the labial sounds in English [pʰ]i[t] and s[p]i[t] differ in aspiration. However, whether the labial will be aspirated is predictable from context; [pʰ] occurs syllable-initially and [p] does not. In this distribution, it is predictable which of the two sounds will occur in a certain environment (complementary distribution), providing no reason to encode the phonetic difference at the phonemic level. Thus, [pʰ] and [p] are allophones of the same phoneme in English, despite their phonetic differences. There are also instances in which two sounds can represent the same phoneme because the meaning of lexical items does not change regardless of which sound is produced. For example, the final sound in cat can be released, i.e. ca[t], or not, i.e. ca[t̚], without
changing the meaning of the word. Because the phonetic difference does not correspond to a difference in meaning, [t] and [t̚] do not need to be distinguished at an underlying level; they are allophones of the same phoneme in this word. Similar to the previous two examples, the final consonants in Russian ma[t] ‘obscene language’ and ma[t̚] ‘mother’ are distinguished by only one feature: palatalization. However, in this instance the phonetic difference between the consonants is not predictable from context and switching one sound for another in these particular words changes the meaning of the lexical item, making it necessary to encode the difference between [t] and [t̚] at the phonemic level. Thus, they are allophones of different phonemes, despite the phonetic similarity between them.

These examples illustrate that two primary criteria used in determining whether sounds are contrastive. Contrastive sounds can change word meaning and are unpredictably distributed. In conjunction with change in meaning, the criterion of predictability, i.e. complementary distribution, is the most commonly used trait to determine where phonemic contrast exists (Dresher 2011:252). This type of traditional phonological analysis plays a primary role in uncovering the relationships between sounds -- it separates sounds into two categories for contrast: those that can distinguish lexical items (not fully predictable and change meaning) and those that cannot (either predictable or do not change meaning).

---

4 Other criteria are sometimes considered such as phonetic similarity, native-speaker intuitions, alternations, borrowed vs. native vocabulary, etc. These criteria are often supplemental to the criterion of (complementary) distribution and used in order to help in unclear or controversial cases.
2.1 Utilization of contrast

While the potential to distinguish lexical items is central to the definition of contrast, it is not clear that potential alone is sufficient in classifying phonological relationships. Phonological theory has long recognized instances where the potential of certain sounds to distinguish words is systematically non-present, i.e. it is neutralized. Neutralization is a common phonological phenomenon where a contrast that is present in certain contexts does not surface in (an)other context(s) because of the phonetic environment. For example, Russian voiced and voiceless obstruents contrast before vowels, e.g. [t]am 'there' vs. [d]am 'lady.GEN.PL', but when /d/ occurs word-finally, it devoices. It is clear from morphophonological alternations that voiced segments are present underlyingly, e.g. mē[d]a 'honey.GEN' ~ mē[t] 'honey.NOM', but they do not surface word-finally because of the phonetic environment⁵. Thus, the systematic contrast between voiced and voiceless obstruents in Russian is not utilized (systematically or otherwise) word-finally⁶.

Neutralization defines those contexts where phonetic factors prevent contrast from being utilized, leaving all other contexts equally available to utilize contrasts in the language. This leads to an expectation that contrasts will be utilized to some degree, if not robustly, in all contexts where neutralization does not prevent them from being utilized. There are, however, contrasts that are only minimally utilized or not utilized altogether. For example, some sounds exist only in recently borrowed vocabulary.

⁵ Whether the devoicing of word-final obstruents in Russian results in a complete neutralization is debated (see, e.g., Shrager (2012) for a recent account).

⁶ Here I use the term "utilized" to mean that the sounds occur in words where they are available to the speaker as distinguishing sounds for that word. This is not to say that they must be the only distinguishing sound(s) in the words, i.e. minimal pairs, for them to be utilized in my terminology.
Defining such sounds as contrastive leads to unsatisfying implications because contrast implies systematicity, something such sounds do not exhibit. In fact, given that contrast is the fundamental device used to determine which sounds are underlying elements of the phonological system, it seems inappropriate to classify sounds as such when they are not systematically utilized in the language. In defining and describing phonological relationships, scholars have recognized the divide between potential to distinguish words and its utilization, but they have had difficulty dealing with it for exactly this reason -- contrast implies a systematicity that not all contrastive sounds exhibit.

2.2 Intermediate relationships

In an attempt to deal with the disjoint between the traditional definition of contrast (potential) and actual sound distributions occurring in real language (utilization), Goldsmith (1995:11) describes three types of phonological relationships that are neither clearly allophonic nor clearly contrastive, paraphrased here.

- "Just-barely contrastive sounds" are in complementary distribution in most contexts but distinct in a limited number of contexts.
- "Not-yet-integrated semi-contrasts" are sounds that, in at least some environments, are in (near) complementary distribution with the exception of a few recent and transparent borrowings.
- "Modest asymmetry cases" are sounds that must clearly be contrastive in the underlying inventory but for which the contrast is neutralized in a variety of contexts.

Goldsmith suggests that these different types of intermediate phonological relationships form a cline of contrast: (fully) contrastive > modest asymmetry case > not-yet-integrated
semi-contrasts > just-barely contrastive > allophones in complementary distribution (12).

On the left side of the cline the phonology plays no role in determining the sounds; the lexicon dictates which sound will occur. On the right side of the cline the opposite is true: the phonology determines which sound will occur based on the context in which it occurs. Despite their differences, Goldsmith notes that the intermediate relationships are "all treated as 'lexical contrasts'" (13). This is because they fit the traditional definition of contrast: in at least one context they have the potential to distinguish words. However, each intermediate relationship differs in the extent to which this potential to distinguish words is actually utilized in the language, leading Goldsmith to suggest that it is inappropriate to give them a single, simplistic classification. In this way, Goldsmith is implicitly arguing that the potential to distinguish words (contrast) is insufficient in defining phonological relationships when taken without regard for how systematically the potential is utilized. Goldsmith is not alone in his uneasiness to use the binary notion of contrast to categorize all phonological relationships. In fact, many scholars avoid using the binary categorization of contrast; instead, scholars hedge in their determination of contrastiveness/phonemicity by creating new terms to describe the sounds that exhibit intermediate relationships. Such terms include 'semi-phonemic', 'quasi-phonemic', 'deep allophone', 'partial contrast', 'fuzzy contrast', 'mushy phoneme', 'near-phoneme', etc. (see Hall 2009:7-8 for a more complete list). Scobbie and Stuart-Smith (2008:106) even argue that "every language has a rump of potential/actual near-phonemes" (emphasis theirs) suggesting that intermediate relationships are not exceptional; rather, they are pervasive in phonology. By intentionally avoiding the classic division between sounds that can and cannot distinguish meaning, each of these scholars recognizes that the potential to
distinguish words, without reference to its utilization, is an insufficient criterion for defining phonological relationships.

2.3 Two dimensions of contrast utilization

The extent to which a contrast is utilized varies on two dimensions: contexts in which a contrast is utilized and number of lexical items that utilize the contrast. Each of these dimensions differentiates utilization of contrast from the traditional definition of contrast as a potential: two sounds are contrastive in a language if they can, in at least one context, distinguish words. In this traditional definition of contrast only one (minimal) pair of words in one context necessitates that the sounds be considered contrastive in a language. Goldsmith's intermediate categories illustrate the divide between the definition of contrast and how contrast is utilized because each of his three categories is intermediate, but for different reasons. Modest asymmetry cases are clearly contrastive in the underlying inventory of the language because the contrast is highly evident in at least one context. However, in these cases the contrast is neutralized in a variety of contexts, problematizing the idea that one context is fully sufficient to define a contrastive relationship. On the other hand, the key trait of just-barely contrastive sounds is that only a limited number of lexical items utilize the contrast, problematizing the idea that one minimal pair is sufficient to define a contrastive relationship. Thus, for a pair of sounds to be highly contrastive, they must exhibit contrast over a wide range of lexical items and in a wide range of contexts. Intermediate relationships differ from this utilization of contrast on one or both dimensions. See Figure 1 for an illustration.
Figure 1. Two dimensions of contrast utilization

Exact placement of the intermediate categories in Figure 1 is somewhat arbitrary because the two dimensions along which these relationships vary are continuous, not categorical. For example, I have placed not-yet-integrated-semi contrasts in approximately the middle of the y-axis, though Goldsmith's description does not specify any limitation on the number of contexts in which such a relationship might appear. The key element of these contrasts is that they occur in a limited number of lexical items. One might also expect different numbers of lexical items to utilize contrasts that exist in a wide variety of contexts. Because one minimal pair in each context is traditionally sufficient to determine that contrast is possible, the extent to which the contrast is utilized in each context becomes somewhat uninformative, leaving it under-investigated or ignored altogether (as will be seen in Chapter 4 below).
2.4 Intermediate phonological relationships: a Russian example

To illustrate in a more concrete way the difficulty presented by less-than-fully utilized contrast, in this section I discuss the distribution of palatalized and non-palatalized velars in Russian.

2.4.1 Velars and context: before consonants and word-finally

Before consonants it is largely the case that only non-palatalized velars occur\(^7\). In previous centuries palatalized velars occurred before palatalized velars as an effect of regressive palatalization. However, this context was one of the first in which regressive palatalization ceased to occur, leaving only non-palatalized velars occurring before consonants in contemporary Russian (Timberlake 2004:62).

Similarly, only non-palatalized velars occur word-finally in Russian. In this particular context, the non-occurrence of palatalized velars is an accidental result of historical developments (see, e.g., Flier 1998 and Padgett 2003b). In other words, while palatalized velars do not occur word-finally, there is no apparent synchronic explanation for their non-occurrence. There is no evidence to suggest that either neutralization or a phonotactic constraint prevents palatalized velars from occurring word-finally. From the perspective of contrast as the potential to distinguish words, this situation is somewhat problematic. If a contrast is not utilized in a context and there is no principled reason that it does not occur, is the potential to distinguish words really present? Neutralization is a mechanism to describe why contrasts are not utilized in some contexts; however, no such

\(^7\) While true in general there are actually two word types in my corpus (see §3.1 below for a description) that, based on orthography, violate this generalization. They are *nix’t* (German for ‘no’) and *offentlix’kajt* (German for ‘public’). Both of these instances are clearly transliterations of German into Russian where the *x’* (Cyrillic *хь*) is an attempt to represent the German sound. Jones and Ward (1969:144) use this German sound as a comparison to help the language learner understand the pronunciation of *[x̑]*. It is not clear whether these two particular items are actually produced as palatalized velars and even if they are, we will set them aside for the current investigation. No such examples are attested in my corpus for <k> or <g>.  

13
mechanism exists to describe contrasts that do not occur, for which there is no clear (synchronic) motivating factor. Thus, it becomes unclear altogether whether contrast really exists in this context. There is no evidence for neutralization, and scholars would balk at describing velars as widespread neutralization word-finally and before consonants. At the same time, there are no examples that illustrate that contrast exists in this context, leaving the question of whether they can distinguish words. In this context a potential contrast is merely not utilized.

2.4.2 Velars and context: velars before vowels

In most instances palatalized velars occur before front vowels and non-palatalized velars occur elsewhere. However, there are a few exceptions where palatalized velars occur before non-front vowels\(^8\), in borrowings such as \textit{barbe}[k\textsuperscript{j}u] ‘barbeque’, \textit{li}[k\textsuperscript{j}o]r ‘liqueur’ and \textit{mani}[k\textsuperscript{j}u]r ‘manicure’, some of which are very old. They likewise appear in foreign names such as [g\textsuperscript{j}o]te ‘Goethe’. Palatalized velars occur in a select number of native lexical items as well. For example, \textit{tkat}’ ‘to weave’ has inflectional forms such as \textit{tl[k\textsuperscript{j}]oš’} ‘you-SG weave’, in which a palatalized velar occurs before a non-front vowel. This also occurs dialectally with a number of velar stem verbs (e.g. \textit{žeč}’ ‘to burn’ and \textit{pec}’ to bake’), though \textit{tkat}’ is the only one accepted in the standard language (Timberlake 2004:59). Similarly there are some grammatical forms, though not universally accepted,

\(^8\) There are also some instances where non-palatalized velars occur before front vowels. Historically all consonants that have both palatalized and non-palatalized allophones (see description of ‘paired’ consonants in §4.1 below) were palatalized before allophones of the phoneme /e/. This constraint has been eroded in several ways, some of which affect velars. Prepositions which consist of a single consonant (e.g. \textit{k} ‘to’), though part of the phonological word, are not palatalized before allophones of /e/, e.g. [k\textsuperscript{c}]tomu ‘to this-DAT’. This also occurs in compounds at the boundary between compounding elements, e.g. [d\textsuperscript{c}v]\textsuperscript{x}tažnyj ‘two-storied’, and in some foreign words (see Antonyuk-Yudina 2010 for details). These instances are counter examples to the near complementary distribution of (non-)palatalized velars; however, I do not focus on them for two reasons: (1) these facts do not only apply to velars, i.e. they are symptomatic of a larger class of sounds, and (2) some of these examples, e.g. palatalization of consonants before allophones of /e/ in foreign words, are not measurable in the resources available for Russian.
such as bere[ɡʃa] ‘protecting’, and names for young animals, such as viloro[ɡʃo]nok ‘young prong-horn’, though there is variation between [ɡʃ] and [z] in young animal names (Jones and Ward 1969:116).

Importantly, the difference between palatalized and non-palatalized velars can distinguish word meaning. This is most evident in a minimal pair, such as [ɡʃo]te ‘Goethe’ vs. [go]te ‘Gotha.DAT (city in Germany)’. Thus, the fact that palatalized and non-palatalized velars occur in the same context, before non-front vowels, and can distinguish word meaning are evidence for the contrastive nature of palatalized velars. However, because many examples are loanwords, not widely accepted, or few in number, scholars have come to various conclusions about the phonemic nature of palatalized velars. For example, Avanesov (1956:170) asserts that [kʃ] [ɡʃ] and [xʃ] are only allophones of /k/, /ɡ/ and /x/ (i.e. not phonemes), based on the fact that they never occur in word-final position. In contrast, Jones and Ward (1969) posit that [kʃ] is phonemic because it occurs before non-front vowels, though only in a few words, "nearly all of [which] are of foreign origin" (114). However, they do not consider [ɡʃ] phonemic because it occurs in very few loans words that are "not in common use" (116). They include no discussion about the possibility of [xʃ] being phonemic, suggesting that such is out of the question. Their analysis is similar to that of Panov (2002:24-25) in that [kʃ] is phonemic in contemporary Russian while [ɡʃ] and [xʃ] are not; however, the nature of Panov's work, as a historical progression of the language, provides an interesting comparison because he claims that even [kʃ] was not phonemic as recently as the beginning of the 20th century (96). Thus, in Panov's view [kʃ] has made an appearance as a phoneme in the last century and he
suggests that eventually [gʰ] may follow, though he does not see the same future for [xʰ] (25).

Despite the different conclusions, all of these arguments share an important similarity: they implicitly recognize that the utilization, not just the potential, of these sounds is a key aspect of determining their (non)contrastive nature. This line of reasoning parallels the many scholars who adopt semi/pseudo/quasi-phonemic terms to describe phonological relationships. What these accounts do not share is a way to quantify the utilization of contrast so that it can be consistently used to describe and investigate phonological relationships, leaving the scholars to somewhat arbitrarily assign more importance to some words/contexts than others.

2.5 Utilization of contrast, approaches and questions

As an alternative to defining phonological relationships by their potential to distinguish words, they should be defined by their potential and how systematically that potential is utilized. Unlike the traditional definition of contrast, such an approach allows the phonologist to capture the intermediate nature of phonological relationships whether the sounds are utilized in few contexts or few lexical items. Additionally, doing so will encode the similarity between non-utilization of contrast and neutralization of contrast without losing the power to describe where sounds can distinguish word meaning.

Hall (2009) introduces a model that uses predictability of distribution to quantify the extent to which a contrast is utilized along a continuous spectrum ranging from (entirely) allophonic to (entirely) contrastive. Unlike Goldsmith's cline, Hall's model does not have a discrete number of categories; rather, it uses tools of information theory to quantify the degree to which a contrast is utilized in any given context by pairs of sounds.
This allows direct comparisons to be made between sounds as a static system as well as at
different points in time on both relevant dimensions of contrast utilization. Importantly,
this model is not only a new application of an existing methodology (though it is that) but
represents a shift in the way contrast is conceptualized. Instead of only questioning
whether sounds can distinguish words in a context, this model measures the extent to
which sounds are utilized to distinguish words in that context. As part of this shift in
view, new questions can be asked. Measuring the utilization of contrast on a continuous
scale allows us to ask if all the 'contrastive' sounds are equally utilized. It provides a
mechanism to quantify how intermediate a relationship is, allows it to be directly
compared to other phonological relationships, and allows it to be compared to the
relationships between the same sounds in different contexts. With these questions come
new implications for the phonological system. For example, is [kʲ] utilized more than [gʲ]
and [xʲ] as Jones and Ward (1969) and Panov (2002) suggest, or are all palatalized velars
utilized similarly in Russian? If the degree to which contrastive sounds are utilized differs
gradiently, are the effects of the utilization, e.g. on perception or production, also
gradient? As the potential to distinguish words emerges, to what extent is the potential
utilized? Does the utilization increase monotonically as a neutralizing process recedes, or
do other aspects of the phonological system affect the path of change? In the following
chapters I ask these questions about three sets of Russian data. I argue that these
questions are important for the phonologist, and questions that cannot be asked (or

---

9 Hall's model includes two ways to measure contrast. One, environment-specific entropy, measures
contrast in a given environment and the other, systemic entropy, includes all environments weighted by the
probability of the environment. In this paper I only use the former to illustrate the effects of gradient
utilization of contrast in specific environments.
answered) without conceptualizing contrast in terms of how it is utilized in a gradient, context-specific way.
Chapter 3: Quantifying utilization of contrast

I have argued that not only the potential to distinguish words but also the utilization of such potential is a crucial aspect of phonological relationships. In this chapter I describe Hall's (2009) Probabilistic Phonological Relationship Model, which uses corpus counts and the information theoretic concepts of probability and entropy to quantify the utilization of contrasts in a gradient way on a context-by-context basis. To do this I will first describe the corpus that is the source of the Russian data in this and subsequent chapters. I will then show how the information theoretic concept of entropy can be used to quantify the use of contrast in a way that is both context-specific and gradient, using data about Russian velars as an example.

3.1 The corpus

Because this study investigates the phonetic character of consonants and the contexts in which they occur, it was crucial to find a corpus consisting of phonetic representation, and not merely orthographic representation. However, phonetically transcribed corpora for Russian are scarce and those that do exist are inadequate because they have a severely limited search function. For this study I thus use a 33.8 million word subset of the Russian National Corpus (RNC) that was algorithmically converted from orthography to phonetic transcription by Robert Daland. I am very grateful to Robert for letting me use this corpus.
was modeled after the British National Corpus and includes a variety of text types, including fiction, newspaper articles and orthographically transcribed speech.

Daland used a lemmatized and part-of-speech-tagged fragment of the RNC corpus that was provided by Serge Sharov (University of Leeds). He wrote a set of rules to derive a phonetic transcription from the orthographic corpus. Russian employs a morphophonemic spelling system in which there is a moderately close correspondence between letters and sounds, but with a number of exceptions. For example, vowel reduction occurs in unstressed syllables and differs depending on the palatalization of the preceding consonant and, for some vowels, the palatalization of the following consonant (see Appendix A for a table of the co-occurrence of vowels, stress and palatalized/non-palatalized consonants). This pattern of vowel reduction is not directly reflected in the orthographic system. However, the phonetic vowel can be determined if word stress and palatalization of surrounding consonants are known. Similarly, regressive voicing assimilation occurs in obstruent clusters, and this is not always represented orthographically (e.g. when a morpheme boundary falls between the consonants in the cluster). Still, given that voicing assimilation applies universally in such clusters, whether a non-final obstruent in a cluster is voiced can be determined from the voicing of the final obstruent in the cluster. By and large, then, the pronunciation of a word is predictable from its orthographic form, even if it is not directly encoded in its spelling.

To determine the phonetic transcription of each token, Daland recovered a phoneme string from the orthography based on standard language pronunciation rules for soft signs, jotated vowels (e.g. <я, ю, и>) and general rules of Russian spelling. He then assigned a stress code to each token by looking up the lemma in Zaliznjak (1977). For
those words that were not in Zaliznjak’s grammar, he made a probabilistic guess about
stress placement. Then, with a derived phoneme string and stress code for each token, he
wrote rules to apply phonological processes such as vowel reduction, word-final
devoicing, voicing assimilation in clusters, etc. See Daland (2009:135-140) for a full
description of the process. Because the corpus was built using a number of computer
algorithms and hand-written rules, it necessarily includes a degree of error, an
unavoidable aspect of corpus work\textsuperscript{11}. Though some details of the data might change with
a cleaner (or larger) corpus, I expect that the major conclusions of the analysis would be
similar.

3.2 Distribution of velars in the Russian corpus

Having described the corpus, we can now use it as a tool to investigate velars in
Russian, our example of an intermediate phonological relationship. Both palatalized and
non-palatalized velars occur before non-front vowels where they can distinguish word
meaning, evidence that they are contrastive. But we might also ask to what extent these
sounds are available as contrastive sounds to the speaker. Do they appear in a large
number of lexical items and how lopsided is their distribution in the context where they
contrast, i.e. before non-front vowels? To answer this question, I searched through the
phonetic corpus and found all individual word types in which a velar occurs before a
non-front vowel\textsuperscript{12}. The data are shown in Table 1.

\textsuperscript{11} Some aspects of the error in the phonetic transcription are more relevant than others for this paper. For
example, all dentals occurring before a palatalized consonant were transcribed as palatalized. As will be
seen below (see §5.3.1) this grossly overgeneralizes the phenomenon of regressive palatalization. This is
one of the reasons I use the orthographic corpus for the data in Chapter 5.

\textsuperscript{12} Here I use word type counts but word token counts, counting each instance of a word type individually,
can also be used in this model. In earlier drafts of this paper I included both type and token counts, but
since the generalizations over the data were largely the same, I chose to present only type counts.
Before non-front vowels 
\[ \{a, o, u, ə, ʌ\} \]

<table>
<thead>
<tr>
<th>Sound</th>
<th>Number of Word Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>[k]</td>
<td>173,747</td>
</tr>
<tr>
<td>[kʲ]</td>
<td>282</td>
</tr>
<tr>
<td>[g]</td>
<td>46,463</td>
</tr>
<tr>
<td>[gʲ]</td>
<td>135</td>
</tr>
<tr>
<td>[x]</td>
<td>25,991</td>
</tr>
<tr>
<td>[xʲ]</td>
<td>38</td>
</tr>
</tbody>
</table>

**Table 1.** Distribution of Russian velars before non-front vowels

It is clear in Table 1 that palatalized velars before non-front vowels are uncommon compared to non-palatalized velars; they do, however, occur. This lopsided nature of the distribution is only trivially important in determining whether there is potential to distinguish words -- the sounds are not fully predictable. The distribution of the sounds is very useful, however, in quantifying the extent to which the contrast is utilized.

### 3.3 Operationalizing distributions: probability

In Hall’s model, sound distributions are operationalized using probability theory. The probability of a sound expresses how often it occurs in a given context as opposed to another sound\(^\text{14}\). For example, before non-front vowels [k] occurs in 173,747 word types

---

\(^{13}\) I do not include [ɨ] here because it occurs as a predictable allophone of the front vowel /i/, as in e.g. б[i]/‘to be’. It also occurs as an allophone of non-front vowels such as /o/, e.g. ё[ɨ]na ‘wife’, in unstressed syllables after non-palatalized consonants that do not have a palatalized counterpart (see §4.1 below). This context makes it irrelevant here for velars and the subsequent investigations in Chapters 4 and 5 below.

\(^{14}\) It is worth noting that Hall’s model restricts the investigation to pairs of sounds, resulting in a Bernoulli distribution where the probability of one sound is always one minus the probability of the other. In any natural language there are likely to be more than two sounds that contrast in any given context; however, it is the relationship between two sounds that illustrates the contrast of a particular (set of) feature(s) and therefore only those observations that include one of the two sounds in question are included for calculating probabilities. Thus, any words that do not contain [k] or [kʲ] before a non-front vowel are irrelevant to the calculation of p(k|A) and p(kʲ|A).
and [k¹] occurs in 282. Thus, if A stands for any following non-front vowel, \( p(k|A) = \frac{173,747}{173,747 + 282} = 0.9983 \). This probability represents the likelihood that [k] will occur, as opposed to [k¹], before a non-front vowel. For the probability of different velars given a following non-front vowel in Russian, see Table 2. (Here, C stands for the consonant in question.) For illustration purposes, Table 2 also includes type counts and probabilities for velars word-finally, which will come into play in the discussion below.

<table>
<thead>
<tr>
<th>Before non-front vowels [a,o,u,ə,ʌ]</th>
<th>Word-final</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>number of word types</td>
</tr>
<tr>
<td>[k]</td>
<td>173,747</td>
</tr>
<tr>
<td>[k¹]</td>
<td>282</td>
</tr>
<tr>
<td>[g]</td>
<td>46,463</td>
</tr>
<tr>
<td>[g¹]</td>
<td>135</td>
</tr>
<tr>
<td>[x]</td>
<td>25,991</td>
</tr>
<tr>
<td>[x¹]</td>
<td>38</td>
</tr>
</tbody>
</table>

Table 2. Probability for Russian velars in two contexts

From Table 2 we can see that probability encodes the extent to which two sounds are predictably distributed -- whether they are fully predictable in a context (velars word-finally) or unpredictable to some degree (velars before non-front vowels).

### 3.4 Uncertainty and utilization of contrast

Table 2 shows the probability of individual sounds in given contexts. Contrast, on the other hand, is necessarily a relationship between (at least two) sounds. Hall employs entropy, a formal measure of uncertainty (Shannon and Weaver 1949), measured in bits, to capture this relationship. To illustrate, let us consider a hypothetical example. Say a

---

¹² Word-final obstruents devoice in Russian. Thus, [g] does not occur word-finally.
speaker of Russian heard a word but, for some reason or another, could not tell if a given sound was palatalized or not. Say the speaker knows the sound was either a [k] or [kʲ] but has to guess which. If the sound were at the end of a word, the speaker would be certain that the sound must be [k] because [kʲ] does not occur word-finally. However, if the sound were before a non-front vowel, the speaker would be less certain which sound occurred. Because of the lopsided distribution in this context the speaker would likely guess [k], but would remain uncertain to some degree. In a distributional system such as this, entropy is the average level of uncertainty induced by the choice between the two sounds. Elements that are completely predictable induce no uncertainty; elements that are mostly predictable induce some, though little, uncertainty; elements that are mostly unpredictable induce larger degrees of uncertainty. For Russian velars, whether a palatalized velar will occur word-finally is fully predictable and therefore induces no uncertainty: entropy = zero. Before non-front vowels, whether a palatalized velar will occur is mostly predictable, though not completely, and therefore a small degree of uncertainty exists.

Recall that (the lack of) predictability is a primary criterion in determining which sounds are contrastive. Not all sounds that are unpredictably distributed, however, are contrastive. So-called free variation is where unpredictably distributed sounds do not affect word meaning. Thus, while Hall's model encodes the unpredictable nature of sounds, the phonologist must determine whether the unpredictability represents contrast or variation. In cases where the sounds clearly distinguish words, the uncertainty induced by the choice between the sounds is an operationalization of the extent to which the contrast between the sounds is utilized. Pairs of sounds that have the potential to distinguish
words may utilize the potential to different extents, resulting in different levels of uncertainty.

3.5 Operationalizing uncertainty: entropy

In Hall's model, where each measurement of entropy is limited to a choice between two sounds, entropy ranges from 0 to 1 bits. Entropy is represented by the Greek letter H and is calculated using the formula \( H = - \sum p_i \log_2 p_i \) where \( p_i \) is the probability of each sound in the chosen context. Using the probabilities derived from Table 2 above we can calculate the entropy induced by the uncertainty between each pair of sounds before non-front vowels and word-finally. Entropy of the distributional system (noted here as K) in which \([k]\) and \([k^\prime]\) are the elements that occur before non-front vowels:

\[
\begin{align*}
p(k\|A) &= 0.9983 \\
p(k^\prime\|A) &= 0.0016 \\
H(K\|A) &= - \sum p_i \log_2 p_i \\
&= -(0.9983 \log_2 0.9983) + (0.0016 \log_2 0.0016) \\
&= -(-0.0023) + (-0.0150) \\
&= 0.0173
\end{align*}
\]

<table>
<thead>
<tr>
<th>Before non-front vowels [a,o,u,ə,ʌ]</th>
<th>Word-final</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of word types</td>
<td>p(C|A)</td>
</tr>
<tr>
<td>[k] 173,747</td>
<td>0.9983</td>
</tr>
<tr>
<td>[k^\prime] 282</td>
<td>0.0016</td>
</tr>
<tr>
<td>[g] 46,463</td>
<td>0.9971</td>
</tr>
<tr>
<td>[g^\prime] 135</td>
<td>0.0029</td>
</tr>
<tr>
<td>[x] 25,991</td>
<td>0.9985</td>
</tr>
<tr>
<td>[x^\prime] 38</td>
<td>0.0014</td>
</tr>
</tbody>
</table>

Table 3. Entropy for Russian velars in two contexts

\[16\] Word-final obstruents devoice in Russian. Thus, \([g]\) does not occur word-finally.
Thus, the choice between [k] and [kʰ] before non-front vowels induces an average of 0.0173 bits of uncertainty. The higher the degree of uncertainty, the higher the entropy. Lower uncertainty results in lower entropy. For example, there is no uncertainty word-finally induced by [k] vs. [kʰ] because only [k] occurs word-finally.

3.6 Discussion: utilization of contrast and Russian velars

Systematically including the utilization of contrast as a factor in defining phonological relationships has several advantages over an approach that defines them only in terms of the potential to distinguish words. Instead of arbitrarily assigning importance to some facts and not others, as has been done in previous analyses of velars in Russian (see §1.2.1 above), this measures the extent to which the contrast for palatalization is utilized in each context.

Instead of classifying sounds as phonemic (or non-phonemic), this account says that velars, on average, induce 0.0205 bits of uncertainty before non-front vowels and 0 bits of uncertainty word-finally (and in other contexts, though these data are not shown). Thus, palatalized velars are marginally utilized along both dimensions -- there is only one context in which they are contrastive and in that context there are a limited number of words that exhibit the contrast. In this way the intermediate nature of the relationship is captured and quantified. Additionally, it is clear that the contrast is not utilized word-finally, even without any evidence of neutralization.

Such an account also allows for a direct comparison between the utilization of contrast exhibited by each pair of velars. In their analyses, both Panov (2002) and Jones and Ward (1969) argue that /kʰ/ is phonemic and the /ɡʲ/ and /xʲ/ are not based on arguments of how extensively /kʰ/ appears in the language. Their classification implies
that one, and not the others, can distinguish meaning. The data in Table 3 (and the minimal pair in 2.1.2 above) present a different picture in that all three pairs of velars have the potential to distinguish words (i.e. entropy greater than zero) and that \( [g^\prime]/[g] \), not \([k^\prime]/[k]\), utilizes the potential to the greatest degree. Only \([k^\prime]\) exists before non-front vowels in forms that are universally accepted in the contemporary standard language and which are incontrovertibly non-foreign, e.g. \( t[k^\prime]o\)š’, and \([k^\prime]\) is present in a larger number of lexical items than \([g^\prime]\), facts that previous scholars may have given prime importance in their analyses. My data suggest that \([g^\prime]/[g]\) utilize the contrast to a greater degree than \([k^\prime]/[k]\) or \([x^\prime]/[x]\), suggesting that if any palatalized velar should be considered phonemic, it should be \( /g^\prime/ \). None of the analyses above have \([x^\prime]\) as phonemic. This seems to line up with the fact that, among the velars, \([x]/[x^\prime]\) induce the smallest amount of uncertainty. This does not, however, suggest that \([x]/[x^\prime]\) cannot distinguish words; rather, it suggests that the contrast between them is utilized least among the velars. This also says nothing about whether this contrast will become more systematically utilized over time, a hypothesis Panov rejects. In Russian, the majority of consonants are paired for palatalization, a trait that velars seem to be gradually acquiring before non-front vowels. Though not sufficient alone, this at least partially suggests that the pressure for the contrast between \([x^\prime]/[x]\) to become more fully utilized over time would be similar to the pressure for other velars, contra Panov’s claim.

For velars in Russian, this model quantifies the limited extent to which contrast is utilized in one context. Instead of arbitrarily assigning importance to some forms and not others, this model provides numerical values that can be compared synchronically, as we have done here, and the same methodology can be used at different points in time to
make a diachronic comparison. As more (or fewer) items that have palatalized velars before non-front vowels are adopted into Russian, the level to which the contrast is utilized will change. Such comparisons can be made across both dimensions of contrast: number of lexical items and number of contexts which exhibit the contrast. Thus, this approach provides a way to measure contrast utilization across both relevant dimensions from both synchronic and diachronic perspectives.
Chapter 4: ‘Paired’ consonants word-finally

In this chapter I will investigate one set of consonants that all have the potential to distinguish words. However, I will show that in (at least) one context, word-finally, the utilization of contrast is not equal for all the consonants and that this utilization of contrast has clear parallels with the perception and production of these consonants in this context, effects that are not easily accounted for when only the potential to distinguish words is investigated.

4.1 'Paired' consonants

Traditionally, Russian consonants are divided into four categories with respect to palatalization: paired non-palatalized consonants, paired palatalized consonants, unpaired non-palatalized consonants and unpaired palatalized consonants (see, for example, Timberlake 2004:57). Consonants are classified as 'paired' when both palatalized and non-palatalized consonants exist that are identical for all other features -- voicing, place and manner of articulation. The paired consonants include labials /p, b, m, v/, dental/alveolar consonants /t, d, s, z, l, ɾ/ velars /k, g, x/ and all their palatalized counterparts. Paired consonants take part in regular morphophonological alternations (e.g., ma[t] 'vulgar language' ~ o ma[t]e 'about vulgar language-LOC.SG'), whereas unpaired consonants do not (e.g. du[ʃ]a '(a/the) soul' ~ o du[ʃ]e 'about (a/the) soul-LOC.SG').

17 Timberlake uses the terms mutable/unmutable which, for these consonants, are synonymous with the terms paired/unpaired. In the Slavic tradition, the terms 'soft/hard' are also often used synonymously with the terms 'palatalized/non-palatalized'.
The unpaired consonants include the palatalized/palatal /ʃ/ /tʃ/ and /j/ and non-palatalized consonants /ts/ /ɡ/ and /z/. See Table 4 for a visual representation.

<table>
<thead>
<tr>
<th>Bilabial</th>
<th>Labiodental</th>
<th>Dental or alveolar</th>
<th>Post-alveolar</th>
<th>Retroflex</th>
<th>Palatal</th>
<th>Velar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plosives</td>
<td>p/p̊</td>
<td>b/b̊</td>
<td>t̪/t̪</td>
<td>d̪/d̪</td>
<td></td>
<td>k/k̊</td>
</tr>
<tr>
<td>Nasals</td>
<td>m/m̊</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trill</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>r/r̊</td>
<td></td>
</tr>
<tr>
<td>Fricatives</td>
<td>f/f̊</td>
<td>v/v̊</td>
<td>s̪/s̪</td>
<td>z̪/z̪</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laterals and Approximants</td>
<td></td>
<td></td>
<td>l/l̊</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affricates</td>
<td></td>
<td></td>
<td>ts</td>
<td>tf̊</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4. An inventory of Russian consonants grouped according to palatalization; white = non-palatalized paired consonant/palatalized paired consonant, grey = unpaired palatalized consonants, dark border = unpaired non-palatalized consonants

Word-finally, all non-velar paired consonants are said to be contrastive for palatalization. Minimal pairs exist for a range of consonant types, including at least dental stops *ma[t]* 'mother' vs. *ma[t̪]* 'vulgar language', dental sonorants *ugo[l]* 'corner' vs. *ugo[l̊]* 'coal', (some) nasals *da[n]* 'given-MASC' vs. *da[n̊]* 'tribute', and (some) labials *goto[f]* 'ready' vs. *goto[f̊]* 'make ready!'. This traditional classification implies that all non-velar paired consonants are utilized similarly word-finally in Russian.

In this chapter I investigate to what extent this is true. Specifically, I ask the following questions: Is the contrast between palatalized and non-palatalized paired consonants utilized similarly word-finally in Russian? If not, are clearly categorical divisions apparent, or is the utilization of contrast gradient? Does the level of utilization affect speakers' perception and production of the sounds in this context?
4.2 Method and data

To investigate these questions I used the phonetic version of the corpus described in §3.1 above. One convenient aspect of paired consonants word-finally is that the mapping between orthographic and phonetic forms is relatively straightforward. While obstruents undergo word-final devoicing, there is no allophonic palatalization of word-final consonants. For paired consonants word-finally, palatalized sounds are followed by an orthographic marker, the *mjakij znak* ‘soft sign’, whereas non-palatalized (paired) sounds are not.

I found all words that contained a paired consonant word-finally. I used the search results to calculate the entropy of each pair of consonants (palatalized and non-palatalized) as described in §3.5 above. As a reminder, entropy (H) reflects the uncertainty given a choice between two sounds. In this context each pair of sounds can distinguish meaning, so the entropy is the extent to which the contrast between them is utilized. In this model entropy ranges from 0 to 1 where 0 reflects no contrast (allophonic distribution) and 1 reflects a maximal utilization of contrast. See Table 5.

---

18 Of course, coarticulation is not restricted by word boundaries in continuous speech. Thus, word-final consonants may be affected by the following word. However, this aspect of word-final pronunciation will be set aside for this study. For an account of (allophonic) palatalization across word boundaries in English, see Bush (2001).

19 Thus, there would be no empirical difference if the data were based on the orthographic version of the corpus, as opposed to the phonetic version.

20 Voiced obstruents are not in these data because they devoice word-finally in Russian.
Table 5. Entropy (based on word types) for paired consonants word-finally

Table 5 has been organized so that the entropy calculated on word types for each set of paired consonants increases from top to bottom. We can immediately see that the utilization of contrast among paired consonants word-finally varies greatly. It is noteworthy that there are not easily identifiable, discrete categories for how contrast is utilized. At one end of the spectrum there is [m]/[mʲ], for which there is almost no uncertainty whatsoever, whereas on the other end of the spectrum [t]/[tʲ] exhibits a level of uncertainty nearly equivalent to the odds of a coin toss. Additionally, it is not the case that a few consonant pairs utilize the contrast robustly while other pairs utilize it to an intermediate degree. Instead, consonant pairs are spread across the possible range of

21 These data do, however, support Timberlake's (2004:60) claim that the contrast between palatalized and non-palatalized consonants is favored most by dentals, labials less, and velars least (of all paired consonants). Velars exhibit zero contrast word-finally, labials exhibit some, dentals exhibit more. Additionally, there seem to be a clear preference for the non-palatalized sound among paired consonants. Only [s]/[sʲ] have a bias for the palatalized sound. This may be a result of the so-called reflexive with past tense verbs. The most common allomorph of the reflexive is -s', which occurs with feminine, neuter and plural past tense verbs. Of the 30 most common forms ending in [sʲ], 24 are the result of the reflexive ending with a past tense verb.
contrast utilization. Thus, the utilization of contrast provides a more complete description of these phonological relationships, all of which have the potential to distinguish words.

4.3 Entropy, perception and production

Various factors affect how speakers perceive sounds. Phonetic context is important because sounds are more difficult to perceive in some contexts than others. Phonological contrast also affects how speakers produce sounds. For example, Boomershine et al. (2008) show that English speakers perceive sounds that are contrastive in English and allophonic in Spanish to be more distinct than Spanish speakers do, and vice versa. They also show that contrastive sounds for which the contrast does not surface, i.e. is neutralized, are more perceptually distinct than non-contrastive sounds but less distinct than sounds whose contrast is not neutralized. This suggests that not only the potential to distinguish words, but number of contexts in which the potential is utilized is an important factor in perception. Recall that the number of contexts in which a contrast is not neutralized represents one dimension along which contrast can be utilized. Here I suggest that the other dimension, the number of lexical items which exhibit a contrast, also affects perception and can even affect production.

Kochetov (2004) used a perception task to investigate how well speakers of Russian and Japanese identified palatalized and non-palatalized consonants in a variety of contexts. Palatalized and non-palatalized sounds have the potential to distinguish lexical items in both languages, providing a situation that illustrates the effects of contrast on perception across language and context. Before vowels, both Russian and Japanese speakers perceived palatalized and non-palatalized dental and labial stops well (mean accuracy of 99% and 96% respectively); however, word-finally Russian speakers
perceived the same consonants better (75%) than Japanese speakers (52%). Importantly, contrast for palatalization is not utilized word-finally in Japanese. This suggests that both phonetic context and phonological contrast affect perception. Speakers of both languages perceived the consonants worse word-finally than before vowels, an effect of phonetic context. However, Russian speakers perceived the consonants better word-finally than Japanese speakers, because the contrast is utilized word-finally in Russian but not in Japanese. These results are consistent with Boomershine et al. in that the utilization of a contrast in a particular context affects the perception of the contrast for speakers. Additionally, Kochetov shows that for Russian speakers, palatalized labials are more difficult to accurately identify (45% accuracy) than palatalized dentals (79% accuracy) word-finally. Interestingly, this perceptual difference parallels the degree of utilization of contrast for palatalization for these consonants shown in Table 5 above; palatalized labials utilize contrast less with non-palatalized labials than palatalized and non-palatalized dentals word-finally. This suggests that the number of words which utilize a contrast affects speakers' perception of sounds in a given context. Effects of low contrast utilization are also notable in aspects of production. Labials are not only perceived badly word-finally, but Timberlake (2004:59) states that there is even a ‘slight tendency’ to produce non-palatalized labials word-finally where a palatalized one is expected. Thus, both dimensions of contrast utilization play an important role in the perception of sounds by speakers. Sounds whose contrast is neutralized in certain contexts are perceived as more similar in those contexts and sounds whose contrast is utilized in a small number of lexical items are not perceived (or even produced in the case of Russian) as well as contrasts that are more fully utilized.
Thus, it seems to be the case that higher utilization of contrast along each dimension leads to greater perceptibility. A definition of contrast that classifies all paired consonants as merely 'contrastive' word-finally loses power in describing and predicting such gradient, context-sensitive effects. All non-velar paired consonants may robustly utilize a contrast before non-front vowels, illustrating that they have the potential to distinguish words in Russian; however, it is the extent to which this potential is utilized word-finally that seems to affect their perception and, to some degree, their production in that context.
Chapter 5: Consonants before palatalized consonants; regressive palatalization

We now turn to one final set of data: consonants before palatalized consonants. Like paired consonants word-finally, this set of data will focus on the utilization of contrast in one particular context. However, this particular data set will illustrate another aspect of contrast utilization -- how it can (potentially) change over time. This situation provides the necessary conditions for such an investigation because it involves a process which affects the degree to which contrast is utilized and which has been declining in Russian over the last century.

Conceptualizing contrast as both the potential to distinguish words and the utilization of that potential is key in asking new questions about this change and its impact on the phonological system. Does contrast ever get utilized as a neutralizing process disappears? And, if so, does the degree of utilization change monotonically or is the direction of the change more complex? If it does not change monotonically, what affects the direction of the change? To address these questions I first review the facts about regressive palatalization. I then present data collected from participants to illustrate the synchronic status of the contrast. I then illustrate the way in which contrast is shifting from neutralization to lack of utilization and discuss the implications of this shift.

22 Here I limit my investigation to certain consonants before palatalized consonants though the majority of the questions are equally relevant for consonants before non-palatalized consonants. For example, consonants exhibit regressive depalatalization, e.g. pja[t] 'five' ~ pja[t]adcat 'fifteen' though this process is less well known and less well studied than regressive palatalization. Before non-palatalized consonants the contrast for palatalization is similarly limited for (non-lateral) dentals where mostly non-palatalized consonants occur before non-palatalized consonants, with a few counterexamples.
5.1 Regressive palatalization: background

When a paired consonant occurs immediately before a palatalized consonant it may be palatalized in Russian\(^{23}\). However, this process, regressive palatalization, does not occur consistently. For example, some speakers palatalize (and all voice) the word-initial /s/ in \([zd]\)elat’ ‘to do’, whereas speakers do not palatalize the /v/ in \([vd]\)elat’ ‘to set’ (Krysin 1974). The likelihood of regressive palatalization seems to be based on a number of factors, including the nature of preceding and following vowels, voicing, stress, morphological position within the phonological word, analogy to other forms, speech rate, register, level of speakers’ education, speakers’ socioeconomic status, speakers’ place of birth and parents’ birthplace (see Drage 1967, Avanesov 1972 and Krysin 1974). However, Drage (1967) argues that only the phonetic similarity (place and manner) of the consonants in the cluster and whether they cross a morphological boundary have any primary importance.

In summarizing these studies, Timberlake (2004:61-62) gives several generalizations for when regressive palatalization is more (or less) likely to occur. He provides two hierarchies to illustrate tendencies for similarity of place and manner. For place of articulation he gives the following (where > can be read as ‘is more likely to assimilate than’): Dental,Dental > Dental,Labial > Labial,Labial > Labial,Dental. This hierarchy illustrates that dentals are more likely as targets of assimilation than labials and the same place of articulation favors assimilation. For manner of articulation Timberlake gives: Fricative,Fricative > Fricative,Stop > Stop,Stop > Stop,Fricative. This hierarchy

\(^{23}\) As one might expect, unpaired non-palatalized consonants are not palatalized before palatalized consonants. Unpaired palatalized consonants can condition palatalization when they appear after a paired consonant. For example in \(l\bar{e}t\check{c}ik\) ‘pilot’ the cluster was pronounced \([f\bar{t}]\) a few decades ago (Avanesov 1972:118).
illustrates that fricatives are more likely to assimilate than stops, and consonants with the same manner of articulation assimilate more often than those that differ in manner of articulation.

5.2 Decline of regressive palatalization and contrast

Another important consideration for regressive palatalization is that the overall likelihood of regressive palatalization has decreased in recent history. As mentioned in Chapter 1, Panov (2002:113) claims that dentals were palatalized before palatalized dentals and labials at the beginning of the 20th century. Thus, for dentals, contrast for palatalization was neutralized in this context because a contrast that was present elsewhere in the language did not surface due to phonetic context. However, as regressive palatalization began to decline, non-palatalized dentals began occurring before palatalized dentals and labials. Thus, the contrast was no longer neutralized, leading Panov (2002:126) to reclassify dentals in this context as having gone from weak (no contrast possible; neutralization) to strong (phonemic contrast possible). This represents the crucial criterion of a traditional definition of contrast -- there is a potential for the sounds to distinguish words. This marks the first step in the possible change in utilization of the contrast. To illustrate this change in how contrast is (not) utilized in this context on a more conceptual scale, let us imagine it progressing from one extreme to another. In Figure 2 this is represented as a progression from Stage 1 to Stage 2 to Stage 3 where in each stage fewer and fewer consonants are regressively palatalized in their conditioning.

24 The laterals present a unique situation here because, though dental, they are not included in Panov's discussion of dentals. Other dentals, even the dental sonorants [n]/[n'] and the flaps [ɾ]/[ɾ'], are included in the discussion about regressive palatalization but [l]/[l'] are not because they exhibited a contrast in the 19th century when other dentals did not (before dentals and labials). I include the laterals in my discussion of contrast for illustrative purposes though it is clear that treating them as targets for regressive palatalization in the same way as other dentals is somewhat misleading. It seems that while it may have be theoretically possible for [l]/[l'] to be regessively palatalized, it is very rare, if it happens at all.
environment. In Stage 1 all consonants regressively palatalize, neutralizing the contrast between palatalized and non-palatalized dentals in this context. In Stage 2 some consonants regressively palatalize, providing the conditions for contrast to be utilized. In Stage 3 no consonants regressively palatalize.

![Figure 2. Decline of regression palatalization over time](image)

As regressive palatalization declines, we expect the contrast for palatalized and non-palatalized dentals before palatalized consonants to become utilized. Initially it might seem that the contrast should emerge and its utilization should monotonically increase as regressive palatalization declines. However, because the contrast is based on the ratio of palatalized and non-palatalized consonants that occur, we must determine what consonants will surface once the process no longer affects them.

5.3 Neutralization and results of its decline

Morphophonological alternations show that at least some targets of regressive palatalization are underlyingly non-palatalized. For example, in *spa[s]* <спас>MASC.PST’ ‘saved-MASC.PST’ ~ *spa[sʰ]*i <спасти> ‘to save’, a non-palatalized consonant [s] occurs when no palatalized consonant follows, suggesting the palatalization of the [sʰ] in *spa[sʰ]*i is an effect of phonetic context. This suggests that the [sʰ] is actually non-palatalized at an underlying level, i.e. /s/, and leads to an expectation that when the neutralizing process declines, the [sʰ] in *spa[sʰ]*i will no longer be palatalized. This suggests that the potential to distinguish words becomes available to be utilized in this context as regressive
palatalization declines. The majority of the targets for regressive palatalization, however, do not have morphophonological alternations that suggest whether a palatalized or non-palatalized consonant will occur in the absence of regressive palatalization.

Here, for practical purposes, I will adopt the stance that the underlying status (palatalized or non-palatalized) of a target consonant is apparent in Russian orthography: the first consonant in a cluster is non-palatalized unless it has an orthographic soft sign directly following it. While we should be wary of arguments connected to orthography, a few pieces of evidence support this stance. First, consonants that are palatalized before non-palatalized consonants, e.g. [l]/[lʲ] in pol`ka <полька> 'polka (dance)' vs. po[lk]a <полка> 'shelf', always have a soft sign after the first consonant. This suggests that consonants are non-palatalized in this context except when orthographically marked.

A more direct argument for this approach is that, as will be seen below (see §5.4.3), the vast majority of (non-orthographically marked) targets for regressive palatalization are produced as non-palatalized consonants by contemporary speakers. This suggests that contemporary speakers consider the majority of targets for regressive palatalization to be non-palatalized at an underlying level. From the perspective of the learner, this is somewhat surprising given that the majority of productions were palatalized in recent history. In the course of this change, speakers must have determined that the underlying nature of these consonants was different than what occurred at the surface. This type of change is consistent with perceptual studies (see, e.g., Ohala 1981) in which speakers factor out phonetic traits that can be attributed to the phonetic context in which the sound occurs. In the case of regressive palatalization, speakers seem to have
determined that the palatalization of the target was an effect of the following palatalized consonant, not an underlying trait of the consonant itself.

This means that the change in the utilization of contrast before palatalized consonants will be an interaction of two primary factors. The first factor is the extent to which regressive palatalization affects words in contemporary Russian. The more words that exhibit regressive palatalization, the smaller the set of words that can utilize the contrast. The second factor is the ratio of targets for regressive palatalization that are underlingly non-palatalized to those that are palatalized. If regressive palatalization were to no longer occur, the contrast would only be utilized to the extent that some consonants are palatalized (those with a soft sign) before palatalized consonants and others are not (those without). Thus, to get a complete picture of regressive palatalization and how it affects the utilization of contrast before palatalized consonants, we must investigate the extent to which it is synchronically active as well as the distribution of consonants with and without soft signs before palatalized consonants.

5.4 Methods: regressive palatalization synchronically

Given that the number of words that exhibit regressive palatalization directly impacts that extent to which the contrast can be utilized, it is important that we first gain a clearer picture of the extent to which regressive palatalization is active in the modern language. Because regressive palatalization is not a phenomenon that can be measured from a corpus (at least not with the corpora available for Russian), I chose to approach this investigation in the laboratory setting. I will discuss how I collected the data and then discuss the implications of the data for contrast before palatalized consonants.
5.4.1 Stimuli

Words were chosen from the same corpus described in §3.1 above. Words were selected which contained one of 32 clusters, 16 of which were target clusters and 16 of which were control clusters. Each target word contained a consonant cluster in which the second consonant was palatalized and the palatalization of the first consonant was indeterminate (and not followed by an orthographic soft sign). Each control word contained a consonant cluster in which both consonants were non-palatalized. The particular clusters that appeared in the stimulus words are shown in Table 6 below. Geminate clusters were not considered (marked as X in Table 6 below). In the table, two cells are marked '*' because Russian obstruents assimilate in voicing to the last consonant in the cluster, making some combinations, e.g. *[pd], phonotactically impossible.

<table>
<thead>
<tr>
<th>2nd cons. → 1st cons. ↓</th>
<th>Target cluster / Control cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>[pʲ] / [p]</td>
<td>X</td>
</tr>
<tr>
<td>[tʲ] / [t]</td>
<td>[pʲt] / [pt]</td>
</tr>
<tr>
<td></td>
<td>X</td>
</tr>
<tr>
<td>[dʲ] / [p]</td>
<td>[pʲn] / [pn]</td>
</tr>
<tr>
<td>[nʲ] / [n]</td>
<td>[pʲl] / [pl]</td>
</tr>
<tr>
<td>[lʲ] / [l]</td>
<td>[pʲp] / [lp]</td>
</tr>
<tr>
<td></td>
<td>[pʲt] / [lt]</td>
</tr>
<tr>
<td></td>
<td>[lʲp] / [lp]</td>
</tr>
</tbody>
</table>

Table 6. Consonant clusters investigated
For each target cluster six lexical items were chosen. For each control cluster four lexical items were chosen (160 lexical items total). In each lexical item the cluster was preceded by a non-front vowel ([a, o, u, ə, ʌ]). Within each cluster type (targets and controls), items were chosen with varying frequencies, and to represent a range of conditions for morphological boundary (whether the cluster crossed a boundary) and stress (whether the stress fell on the vowel immediately preceding the cluster or elsewhere).

5.4.2 Participants and recording

Five female native Russian speakers digitally recorded the stimuli in a recording studio in Hagerty Hall on The Ohio State University Campus. Participants varied in age from 22 to 50 years old and all had some level of post-secondary education. All participants were speakers of the standard language from various cities in the former USSR who had been living in the U.S. from 3 to 16 years. They were told the study was about Russian pronunciation and were instructed to read each word when it appeared. The words were presented one-by-one on a computer screen in size 44 Arial font. Each word was on the screen for 4 seconds after which it was directly followed by the next word. Participants were not tested to see if they knew/recognized all the words, but conversations with participants after their recordings suggested that they knew all the words. The words were presented in a random order, with the same order being used for all participants. All participants recorded the same list of stimuli, pronouncing each

---

25 More tokens were selected for the target group than the control group in an attempt to tease out what factors most affect the likelihood of regressive palatalization. However, these factors proved to be largely unimportant for this thesis, resulting in an accidentally unequal number of tokens between the controls and targets.

26 Participants lived in the following cities: Kaliningrad (Russia), Kirovohrad (Ukraine), Liozna (Belarus), Moscow (Russia), Rostov-on-Don (Russia), Taganrog (Russia), Tver' (Russia), Volgograd (Russia).
stimulus once. Half-way through the recording for each participant, a short break was
given during which the participants filled out a background questionnaire. Participants
wore headphones (which allowed them to hear themselves) and spoke into a free standing
Electro-Voice RE20 microphone. All recordings were made at a sampling rate of 44.1
kHz with no high or low pass filtering.

5.4.3 Determining palatalization in consonant clusters with acoustic data

Palatalization for consonants is acoustically apparent in two primary ways: the
second formant transition in the preceding vowel and the release of the consonant.
Because the consonants in question were followed by another consonant, the consonants’
release was difficult to analyze, and often non-existent. Thus, the relevant acoustic
correlate for palatalization for consonants before other consonants is the second formant
transition in the preceding vowel.

Boundaries were marked at the beginning and end of the vowel preceding the
cluster in question for all 800 tokens (five speakers, 160 items per speaker). Second
formant measurements were collected for each token at the midpoint of the vowel and 3
milliseconds before the end of the vowel. These were collected using an automated script
that utilized the built-in ‘get formant’ function in Praat. Initial measurements were
checked for irregularities. Errors were fixed by adjusting the input settings for the get
formant function on a case-by-case basis (e.g. maximum formant (Hz), number of
formats, etc.), rerunning the automated script on the token, and then manually verifying
the result.
Once F2 values were retrieved for the mid-point and end of the vowel for each token, the mid-point value was subtracted from the end value to reflect the change in F2 in the vowel. These values were compared across the target and control groups.

Timberlake (2004:61) states that “contemporary speakers have rather less - if any - palatalization than was reported [in earlier studies],” suggesting that regressive palatalization may no longer exist at all in modern Russian. A first question we might ask is if this is the case.

![Figure 3. Comparison of change in F2 in control and targets clusters](image)

In Figure 3 it is visually apparent that the majority of the targets and controls fall within the same range of change in F2. The targets have more productions that are outliers, and the change in F2 of these outliers extends much higher than any productions in the control group. These observations suggest two things: (1) regressive palatalization
does not affect the majority of productions as it once did and (2) some productions, the outliers, still seem to be palatalized as an effect of regressive palatalization.

Another question to address in these data is the extent to which the same lexical items exhibited a large change in F2 (the correlate of palatalization) across speakers. Only 63 productions in the target group (13.13%) have a greater change in F2 than two standard deviations above the mean of the control group (hereafter simply called 'outliers'). Among these 63 productions, 23 lexical items were represented. To investigate the extent to which the productions of these lexical items varied across speakers, I totaled the number of productions that were among the outliers and those that were not for each lexical item. See Table 7.
In Table 7 only five words are produced with a high change in F2 across all speakers. The remaining items exhibit variation in that some speakers’ productions are consistent with regressive palatalization, whereas others are not. This suggests that a select few lexical items still exhibit regressive palatalization consistently across speakers, but that the majority of the outliers are in some state of variation among speakers.

Another question we can address is to what extent the effects of regressive palatalization were similar for each target consonant. To do this I separated the targets and controls by consonant. See Figure 4.

### Table 7. Words with a high change in F2

<table>
<thead>
<tr>
<th>Transliteration</th>
<th>Gloss</th>
<th>Outlier productions</th>
<th>Non-outlier productions</th>
</tr>
</thead>
<tbody>
<tr>
<td>polsotnij</td>
<td>fifty</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>rabotnij</td>
<td>worker</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>segodnija</td>
<td>today</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>subbotnij</td>
<td>(soviet) service day</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>ugodlivyj</td>
<td>obsequious</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>troekratnik</td>
<td>three-timer</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>garantija</td>
<td>guarantee</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>golodnjak</td>
<td>hungry person</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>modnučij</td>
<td>(very) fashionable</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>podlinnyj</td>
<td>genuine</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>bondevik</td>
<td>Bondevik (name)</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>citatičestvo</td>
<td>quotation-mongering</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>gramotnik</td>
<td>literate person</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>komandir</td>
<td>commander</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>postletat’</td>
<td>to fly away</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>rudnik</td>
<td>(a) mine</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>vonlijarskij</td>
<td>Vonlarskij (name)</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>xlopotlivyj</td>
<td>troublesome/fussy</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>balander</td>
<td>food giver (in prison)</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>korrespondent</td>
<td>correspondent</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>otičat’</td>
<td>to distinguish</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>podnimat’</td>
<td>to lift</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>skandinavskij</td>
<td>Scandinavian</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>

In Table 7 only five words are produced with a high change in F2 across all speakers. The remaining items exhibit variation in that some speakers’ productions are consistent with regressive palatalization, whereas others are not. This suggests that a select few lexical items still exhibit regressive palatalization consistently across speakers, but that the majority of the outliers are in some state of variation among speakers.

Another question we can address is to what extent the effects of regressive palatalization were similar for each target consonant. To do this I separated the targets and controls by consonant. See Figure 4.
Previous accounts (such as Panov’s) do not include [l]/[lʲ] with other dentals because [l] is said not to regressively palatalize. These data supported this claim because the target and control group for [l] are not statistically different based on a Welch two sample t-test (p = 0.764; alpha value of 0.01 with Bonferroni correction). The control and target group for [p] do not differ significantly (p = 0.030), suggesting that labials no longer regressively palatalize as they once did. The target and control groups of the remaining three consonant pairs, however, do differ statistically ([d]/[dʲ], p = 0.0034; [t]/[tʲ], p < 0.0001; [n]/[nʲ], p < 0.0001). Thus, regressive palatalization still affects some
clusters in which the target is a (non-lateral) dental. These observations conform to Timberlake's generalization that dentals are more likely targets than labials for regressive palatalization, suggesting that this preference remains even though the scope of regressive palatalization is heavily limited.

5.5 Regressive palatalization and contrast

I now return to the main question of how regressive palatalization and its decline can potentially affect the utilization of contrast that exists for consonants before palatalized consonants. Given that regressive palatalization is very limited in contemporary Russian, the ratio of underlyingly palatalized to non-palatalized consonants before palatalized consonants will play an important role in the utilization of contrast in contemporary Russian. To approach this, I again turn to corpus data.

5.6 Methods 2: regressive palatalization diachronically

For this investigation I use the orthographic version of the corpus described in §3.1 above. The search process included finding all words that had a two-consonant cluster in which the first consonant was an orthographic 'd' <д>, 't' <т>, 'n' <н>, 'l' <л> or 'p' <п> and the second consonant was an orthographic 'd' <д>, 't' <т>, 'n' <н>, 'l' <л> or 'p' <п> followed by a jotated vowel or soft sign (i.e. was phonetically palatalized). Words were then counted in two groups: those with a soft sign after the first consonant, e.g. *oden'te* <оденьте> 'put on! (of clothing)', and those without, e.g. *santimetr* <сантиметр> 'centimeter'. The first consonant was classified as underlyingly non-palatalized unless it was followed by a soft sign, for reasons discussed in §5.3 above. All remaining calculations (probabilities, entropy) were calculated in the same way as for the data sets in Chapter 3 and Chapter 4.
5.7 Change of contrast as regressive palatalization declines

Here I present data to show how the distribution of palatalized and non-palatalized consonants before palatalized consonants would change from one extreme, in which all clusters undergo regressive palatalization (Stage 1 in Figure 2 above), to the other extreme, in which no clusters undergo regressive palatalization (Stage 3 in Figure 2 above). This discussion thus constitutes a hypothetical illustration of the way in which the distribution changed in this environment as regressive palatalization declined. Table 8 also illustrates the change in entropy as the percentage of word types that regressive palatalize declines. I include columns where fifty and ten percent of word types regressively palatalize as two hypothetical points during the change.

<table>
<thead>
<tr>
<th>Number of word types</th>
<th>% of word types reggressively palatalize</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C'C'</td>
</tr>
<tr>
<td>(palatalized)</td>
<td></td>
</tr>
<tr>
<td>[d] vs. [d']</td>
<td>0</td>
</tr>
<tr>
<td>[t] vs. [t']</td>
<td>17</td>
</tr>
<tr>
<td>[p] vs. [p']</td>
<td>7</td>
</tr>
<tr>
<td>[n] vs. [n']</td>
<td>71</td>
</tr>
<tr>
<td>[l] vs. [l']</td>
<td>3923</td>
</tr>
</tbody>
</table>

**Table 8.** Change in entropy as percent of word types that reggressively palatalize changes for five consonants

Table 8 illustrates how the (un)predictability of distribution shifts for palatalized and non-palatalized consonants before palatalized consonants as regressive palatalization declines. In §2.1 above I described two types of phonological relationships that are unpredictably distributed: contrastive sounds and sounds in free variation. Given that Hall’s model is based objectively on the distribution of sounds, it becomes the task of the phonologist to determine whether unpredictably distributed sounds, like those in Table 8, can actually distinguish meaning, i.e. are contrastive, or not, i.e. are in free variation,
which is a problematic task for these data. The clearest evidence that two sounds can
distinguish meaning is a minimal pair. One near minimal pair for palatalized and non-
palatalized dentals before palatalized dentals that occurs in my corpus is ra[n̛lj]e
раньт'е 'hurt/wound.IMP.PL' vs. ra[nlj]e <рантье> 'rentier/investor'.

These words illustrate that when regressive palatalization does not occur, a non-palatalized dental can
contrast with a palatalized dental to distinguish meaning. This does not, however, entail
that all targets for regressive palatalization are able to distinguish meaning in
contemporary Russian. The data in Table 7 (see §5.4.3 above) clearly show that some
lexical items exhibit variation across speakers, suggesting that the target for regressive
palatalization in these words does not have the potential to change word meaning.
Together these two pieces of evidence are problematic for the analytical procedure of
determining whether contrast exists in this context. Given that our discussion of the
change in utilization of the contrast is largely hypothetical, it will suffice here to say that
at least some of the unpredictable nature of the distribution at each point along the change
in Table 8 is due to variation, not contrast. It is worth noting that the items which exhibit
variation comprise only a small percent of all the words investigated, at least in
contemporary Russian. Of the 96 lexical items in the target group (see §5.2.1 above), 18
of them (18.75%) had productions in both the outlier and non-outlier groups, i.e.
exhibited variation. The majority of the words were produced consistently across all
speakers. Given that the majority of words were produced consistently as either
palatalized or non-palatalized consonants, it is plausible to suggest that the nature of the
target consonant could be utilized to distinguish meaning for these words. Thus, most of

---

27 Minimal pairs also exist when the dental or labial is followed by a velar, e.g. re[dk]i 'rare.PL.' vs. re[dk]i
 '(black) radish'.

51
the uncertainty represented by the entropy calculations in Table 8 is a result of contrast, with a small portion of the unpredictability being a result of variation.

Returning to the data in Table 8 we can see that stops (dental and labial) have very few word types in which the first consonant is underlingly palatalized (CjCj in Table 8).28 In contrast, the liquids [l]/[l̩] have a significant number of word types that have an underlingly palatalized first consonant, though there are more with an underlingly non-palatalized first consonant. Because contrast is based on the ratio between palatalized and non-palatalized consonants, the lopsided distribution of the dentals affects the extent to which the contrast is utilized as regressive palatalization declines. Importantly, the change in entropy does not shift equally for all consonants as the percentage of word types that regresses palatalize decreases. For [d]/[d̩], entropy reaches its maximum (H = 1) when half of the word types regresses palatalize. This is because all word types with [d]/[d̩] are targets for regressive palatalization. When half of these word types regresses palatalize, half of the word types are palatalized and half are non-palatalized, leading to a maximal level of uncertainty. After this point entropy begins to fall until it disappears altogether. For [l]/[l̩], entropy changes along a very different path. Because many word types are palatalized whether or not regressive palatalization occurs, regressive palatalization has a less drastic effect on the level of contrast. Entropy rises to its maximum just before the 10 percent mark in Table 8 and

---

28 It is interesting that the vast majority of words that have a CjCj cluster are imperatives. Verbal stems that are stressed and end in a single consonant have a soft sign in the singular imperative and a soft sign followed by -te in the plural, e.g. *otv* CjCj *otv* te <otvetete> 'answer!', vsy CjCj *vs* te <vstane> 'pour (into)!'; vsj CjCj *vst* te <vstane> 'stand!'. This may suggest that in Stage 3 the contrast has become mostly limited to marking a particular inflectional form, rather than being able to distinguish words with different basic meanings.
then declines slightly until no clusters regressively palatalize. This change is visually represented for two sets of consonants in Figure 5.

![Figure 5. Illustration of change in entropy during the decline of regressive palatalization](image)

The case for [l]/[l̄] seems to match the naive expectation that once contrast is possible, it becomes increasingly utilized as regressive palatalization declines. As regressive palatalization declines the contrast is utilized more and more until it peaks just before Stage 3. The change in entropy for [d]/[d̄], however, is very different. Looking at the beginning and end of the change for [d]/[d̄], the extent to which the contrast is utilized is identical; however, only [d̄] existed in Stage 1 whereas only [d] exists in Stage 3. For [d]/[d̄] the decline of regressive palatalization first introduced a contrast and then removed it; Stage 1 represented a neutralization of contrast and Stage 3 represents a non-occurrence of contrast. These two stages share distributional similarities, i.e. full predictability, an observation that is unnoticed altogether if contrast is only investigated as the potential to distinguish words and not the utilization of the potential.
This observation leads to an interesting consequence for the change in the degree to which the contrast between palatalized and non-palatalized dentals has been utilized. Initially, the contrast was increasingly utilized along both dimensions of contrast (see Figure 1 in §2.4 above); the contrast became possible in a new context and new lexical items utilized the contrast. The utilization of the contrast continued to grow in terms of lexical items until Stage 2. By Stage 3, however, the utilization of the contrast for [d]/[dʲ] reverted to being exactly the same as it was in Stage 1: not a single pair of lexical items exhibit the contrast before palatalized consonants. And, the utilization of contrast for other (non-lateral) dentals is only trivially different than Stage 1: a few select lexical items exhibit the contrast in this context. Thus, given that our predictions for how contrast might affect the phonological system (e.g. differences in perception or production) are based on the degree to which a contrast is utilized in a given context, we expect largely the same effects at Stage 1 and Stage 3, a rather counterintuitive prediction for an analysis that treats the potential to distinguish words as the critical element of contrastive sounds. Thus, in this instance, the utilization of contrast is more useful than contrast itself in defining the consequences of the decline of regressive palatalization.
Chapter 6: Conclusion

Despite its ubiquity in phonology, the use of contrast as merely the potential to distinguish words continues to be a problematic notion. In particular, it fails to adequately classify and describe intermediate phonological relationships and their effects on the phonological system, whether in perception/production of the synchronic system or predictions/results of diachronic changes. In this thesis I suggest that not only contrast, but its utilization, are the key traits necessary to adequately define phonological relationships.

Using Hall's (2009) Probabilistic Phonological Relationship Model, I showed how the utilization of contrast can be measured, instead of marginalized, for intermediate phonological relationships, such as palatalized velars. Velars exhibit a limited degree of contrast and in only one context in Russian. My data suggest that all three palatalized velars, \([k^l, g^l, x^l]\), minimally utilize this contrast before non-front vowels despite previous accounts that suggest only \([k^l]\) is utilized sufficiently to be considered contrastive. I also showed that paired consonants exhibit a wide range of contrast for palatalization word-finally in Russian. Some consonants have nearly equal distributions resulting in high uncertainty, e.g. \([t]/[t^l]\), while others are heavily lopsided distributions resulting in low uncertainty, e.g. \([m]/[m^l]\). This uncertainty reflects the gradient degree to which contrasts are utilized by palatalized and non-palatalized consonants word-finally. The perception
and production of these consonants is affected differently depending on how robustly the contrast is utilized in this context. Low contrast corresponds to low perceptual accuracy and even a tendency not to produce the contrast. This reinforces the necessity of conceptualizing phonological relationships by how they utilize contrast, not simply whether contrast is possible or not. I also illustrated the way in which contrast utilization can change over time using the decline of regressive palatalization as an example. I presented new empirical data concerning the synchronic status of regressive palatalization in Russian: it exists but is highly limited in scope. As regressive palatalization declines in Russian, the degree to which consonants utilize contrast shifts. The direction of the shift crucially depends on the balance between underlyingly palatalized and non-palatalized consonants in that context. For dentals, uncertainty peaked and has long been decreasing as regressive palatalization affects fewer and fewer lexical items. Thus, despite the decline of regressive palatalization, dentals seem fated to utilize contrast very little, if at all, before (palatalized) consonants. They have shifted from the neutralization of contrast to a non-utilized contrast, both of which have similar effects for speakers.

In this thesis I have used Russian data to argue that the utilization of contrast is a key trait in defining phonological relationships. Utilization of contrast must be measured in a context-specific and gradient way to account for possible effects on the system, and to make predictions for how contrast might change over time. Not surprisingly, this invites additional questions. The potential to distinguish words has long been used as the crucial criterion for defining phonological relationships. However, I have argued that the utilization of this potential is also necessary to adequately deal with various phonological
phenomena. If the utilization of contrast is a crucial element of defining phonological relationships, is it sufficient? In most instances, the utilization of contrast produces all the distinctions that contrast itself provides as well as additional gradient, context-specific information. One instance where the two diverge is in the distinction between neutralization and non-occurrence of contrasts. From a utilization perspective, these two situations are identical in that contrast is not utilized in either one. From a potential perspective, they differ in that contrast is impossible where neutralization occurs but possible though not utilized in instances of non-occurrence. The important question becomes whether or not these two phenomena should be distinguished. It seems that both lead to less perceptual distinctness for speakers but it is unclear if the effects differ in any significant way. It does seem apparent that speakers can recognize phonetic context as a conditioning factor, suggesting that the difference between neutralization and non-occurrence could be available to the speaker. If this were the case, we would expect speakers to be more willing to allow contrast to be utilized in contexts where it does not occur, e.g. velars word-finally or dentals before palatalized consonant, than in cases where it is neutralized, e.g. word-final devoicing of obstruents. Though speculative, this would suggest that the potential to distinguish words remains a crucial aspect of phonological relationships even when the utilization of such potential is taken into account.
References


Appendix A: Overview of vowel phones and their co-occurrence with stress and (non-)palatalized consonants

<table>
<thead>
<tr>
<th>Phone</th>
<th>Notes</th>
<th>Phone</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>[a]</td>
<td>NA</td>
<td>[o]</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[o]</td>
<td>Between palatalized Cs</td>
</tr>
<tr>
<td>[e]</td>
<td>Heavily restricted</td>
<td>[e]</td>
<td>Before non-palatalized Cs</td>
</tr>
<tr>
<td>[e]</td>
<td>Heavily restricted</td>
<td>[e]</td>
<td></td>
</tr>
<tr>
<td>[u]</td>
<td>[u]</td>
<td>/u/</td>
<td></td>
</tr>
<tr>
<td>[i]</td>
<td>NA</td>
<td>/i/</td>
<td></td>
</tr>
<tr>
<td>[i]</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In unstressed syllables

<table>
<thead>
<tr>
<th>Phone</th>
<th>Notes</th>
<th>Phone</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This table does not include word-initial contexts because they are not important for palatalization of consonants. For vowel phonemes /a, o, u, e/ the same phones that occur after non-palatalized consonants occur word-initially. For /i/ only [i] occurs word-initially.

Most commonly Russian is analyzed as having five vowels phonemes following the Moscow school of thought. In contrast, the Leningrad school of thought posits six vowel phonemes in which /i/ (in the five phoneme system) is split into /i/ and /ɨ/. The difference is based primarily on the fact that suffixes beginning with [i] induce morphophonological alternation in the final base consonant instead of vocalic alternation being conditioned by the consonant, as is the case in other contexts (Timberlake 2004:41). Such an instance of nearly full predictability is central to the main argument of this paper; however, given the focus on consonants in particular, the five vowel system is adopted for simplicity.

Historically these combinations did not occur. This constraint has been eroded in a few ways. Prepositions which consist of a single consonant (e.g. в 'in/into'), though part of the phonological word, are not palatalized before allophones of /e/, e.g. the [v] in [ve]tom ‘in this-LOC’. This also occurs in compounds, e.g. дva[xe]tažnyj ‘two-storied’ and prefixes, e.g. [se]konomi’ ‘to save’. Many foreign words also exist in which non-palatalized consonants occur before allophones of /e/ (see Antonyuk-Yudina 2010 for details).
<table>
<thead>
<tr>
<th>Phone</th>
<th>Notes</th>
<th>Phone</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ʌ]</td>
<td>Pretonic</td>
<td>NA</td>
<td>/a/ and /o/</td>
</tr>
<tr>
<td>[ə]</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[i]</td>
<td>After unpaired non-palatalized Cs</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>[i]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[i]</td>
<td>After unpaired non-palatalized Cs</td>
<td>NA</td>
<td>/e/</td>
</tr>
<tr>
<td>NA</td>
<td>[i]</td>
<td></td>
<td>Sometimes [i]</td>
</tr>
<tr>
<td>NA</td>
<td>[ʊ]</td>
<td></td>
<td>Between palatalized Cs /u/</td>
</tr>
<tr>
<td>[u]</td>
<td>[u]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[i]</td>
<td>NA</td>
<td></td>
<td>/i/</td>
</tr>
<tr>
<td>NA</td>
<td>[i]</td>
<td></td>
<td>Sometimes [i]</td>
</tr>
</tbody>
</table>