



Markedness and Sound Change in OT

Item Type	Article
Authors	Zubritskaya, Katya
Download date	2026-04-15 14:04:36
Link to Item	https://hdl.handle.net/20.500.14394/36818

Markedness and Sound Change in OT

Katya Zubritskaya

University of Pennsylvania

1. Introduction

Over the last few years quantitative examination of on-going sound changes has been the almost exclusive domain of sociolinguistics, remaining outside the purview of mainstream phonological theory. Almost all analyses of sound changes in formal phonology has concerned the reexamination of absolute sound changes (discrete non-variable phenomena), changes that involve reformulating phonological rules or rule reordering. When the actual operation, or mechanism, of sound change are examined quantitatively, however, some of the assumptions of formal phonology are not readily applicable. Since all formal phonological theories are designed to account for categorical phenomena, there are few if any provisions made for the analysis of variable processes. This is one of the reasons why sociolinguists have, for their part, rarely attempted to incorporate the developments of modern phonology into their examination of language change. In this paper I will examine two on-going sound changes in Russian within the formal framework of Optimality theory.¹ This theory offers, I believe, new possibilities for modeling the mechanism of sound change, possibilities absent in other formal theories of phonology. On the other hand, I will demonstrate that the examination of variable data poses a significantly difficult set of problems for current OT formalizing.

The paper has been ordered as follows. In the first part I will focus on the directionality of two sound changes in Russian with respect to various segmental and

¹ I express my thanks to Mark Liberman for his help in the preparation of this paper, and to Elisabeth Selkirk, William Labov and Anthony Kroch for their suggestions and discussion. Errors and shortcomings remain my own.

morphological conditioning environments in order to show that OT provides meaningful explanations for the role of markedness relations in phonological change. While the role of markedness and naturalness in the mechanism of sound change has been the subject of much discussion in linguistic literature, these notions have not been directly dealt with in recent formal theories of phonology (with the notable exceptions of Underspecification and Grounded phonology). Even when markedness is addressed in particular formal theories, the formalisms are designed to account for specific categorical phenomena (such as, for example, the exceptional behavior of coronal segments as a direct reflection of feature geometric representation in Underspecification theory). These formalisms are not equipped to deal with gradient markedness (such as the fact that the Coronal feature may be present in the representation, but is less "marked" than the Labial feature). With the advent of the Optimality Theory, markedness and naturalness relations has once again become a focus of attention in formal phonology, since OT has directly formalized these relations as universal, implicationaly-ordered families of constraints. I will show that the universality and implicational gradience of constraints in OT provide an account of the role of markedness in the directionality of sound changes. This paper will explore the suggestion that whole families of functionally similar constraints may be involved in the process of a sound change so that a change operates as a gradual weakening or a gradual strengthening of a particular constraint family in a grammar. Most importantly, the relations of markedness between individual constraints in a family predict and explain the directionality of a sound change and provide an account for the role of naturalness in the dynamics of a linguistic system. Within this framework a sound change is governed by the same regularities as other categorical phenomena in Russian and cross-linguistically.

In the second part of this paper, I will address a more difficult set of problems. It is a fundamental premise of OT that constraint domination is absolute, and constraint ranking determines the choice of only one optimal candidate from the possible candidates in a set for a particular input form. In this respect OT is modeled like any other formal theory of phonology: it accounts for categorical phenomena only. Since a sound change ultimately involves two distinct rankings (formally grammars in OT) it must be analyzed as constraint reranking. Correspondingly, the most obvious account of variation during the period of a sound change – an account that requires no modification of OT formalisms – is a grammar competition model. In this model, however, the inclusion of all sources of gradience is possible only at the cost of adding possible competing grammars. I will contend that, in the case of Russian, such inclusion of all relevant constraint families in the operation of a sound change would require an absurdly large number of competing grammars. Hence the predictive power of universal implicational ranking within constraint families comes at a considerable cost in a grammar competition model. Secondly, since all constraints are discretely ranked, modeling the sound change as constraint reranking does not provide an easy account of the observed temporal directionality of sound changes. The direct incorporation of markedness relations into a mechanism of a sound change closely resembles the implicational models of sound change proposed in the 70s, most notably by Bailey (wave model of phonological change). Bailey (1973) used the markedness relations to suggest that the more marked environment will be affected in the course of an incoming sound change only after the sound change has been completed in the less marked environment. Kroch (1989), however, in his examination of syntactic change has shown that newer forms advance at the same rate in all relevant environments and the degree to which a particular environment favors or disfavors a new form is constant through time. While phonological change may well turn out to be unlike the syntactic change in some respects, the examination of variable phonological data has consistently shown that sound changes similarly affect all environments simultaneously for most of their duration. While OT makes no claims about the possible mechanisms of a sound change, it is not obvious how the model of constraint reranking can rule out the situation suggested by Bailey. The fact that in a sound change different environments are seemingly affected simultaneously

but at different rates can receive an explanation only as the probability distribution among competing grammars, where grammar competition in one environment implies grammar competition in another, functionally similar environment. Lastly, the model of grammar competition does not provide any easy explanations for the observed frequencies of data distribution, since any frequencies at all can be modeled in principle given several competing grammars. In other words, since the choice of a particular outcome simply means a choice of a particular grammar, any frequencies of data distribution are possible in this model and the internal structure of grammars does not explain anything about the actual observed frequencies.

To summarize: I will show that the mechanism of the sound change quite clearly involves a state of competition between a single constraint with a whole constraint family, where the markedness relations between individual constraints determine the directionality of a sound change in so far as they directly predict which environments will favor or disfavor newer forms. The preservation of categorical ranking and discreteness of constraints in OT grammar, however, necessitates a costly grammar competition model, which provides no insight into the observed patterns of variable data distribution. To provide a solution to this problem, I will explore alternative elaborations of OT formulations, such as the notion of constraint “tie” (noncategorical ranking of competing constraints).

This paper is based on the results of a 1974 sociolinguistic study undertaken in Russia by linguists at the Russian Language Institute. It is the largest available corpus of data on variable phenomena in Standard Russian, and involves a statistical examination of the pronunciation of identical word lists by over 2000 people of varying ages. Additional data is supplied from two smaller studies, also undertaken by Russian linguists, which examine the operation of two sound changes in the speech of about 60 speakers.

2. The Directionality of Sound Change and Markedness

2.1. Palatalization loss

I will begin with an examination of palatalization assimilation loss in consonant clusters in Modern Russian. This change is one of the most striking, as well as one of the most complicated, recent developments in Russian. My analysis of palatalization assimilation is based on the following theoretical assumptions about the feature geometric representation of secondary palatalization:

- a) Palatalization is a secondary coronal feature (Clements 1990, Hume 1992).
- b) Place features are directly parsed by a root node or other place features: there are no intervening C-place, V-place nodes.

At the earliest attested stage, in the beginning of the 20th century, Standard Russian had a regular rule of secondary palatalization assimilation in some nonhomorganic clusters:

- (1) coronal/labial: z'v'er' 'beast'
- coronal/dorsal: s'k'inut' 'throw off'
- labial/dorsal: lap'k'i 'paws'

(Palatalization was also obligatory in homorganic clusters, but the loss of assimilation in that environment is governed by different regularities, which I will not discuss here).

It is clear from the examples I have provided above that the primary place of articulation of consonants in a cluster was the major factor in the operation of assimilation. The relevant descriptive generalizations about the operation of assimilation can be formulated as follows:

- a) If the target is a Coronal, it assimilates to following labial and dorsal consonants
- b) If the target is a Labial, it assimilates only if the trigger is a Dorsal.
- c) If the target is a Dorsal, there is no assimilation.

How can one account for these generalizations in the framework of OT? I suggest that the notions of gradient place markedness, formally captured in OT as a universal markedness scale, are crucial in the characterization of this phenomenon. Following P&S (1993), I assume that grammars generally militate against place feature associations: *PLACE. A prohibition on place features, however, is subject to gradient violations, so that coronal feature is less marked than other places of articulation: *PL LAB » *PL COR. Similarly, the typological asymmetry in the existence of doubly articulated segments is encoded in the dominance hierarchy on consonantal place features parsed under the same root node²:

- (2) *COMPLEX PLACE :
- | | | | | |
|-------|---|------|---|------|
| *DORS | » | *LAB | » | *COR |
| | | | | |
| COR | | COR | | COR |

The role of this hierarchy in Russian phonology is motivated not only by the facts of variable palatalization assimilation, but by a wide array of other categorical phonological phenomena. For example, there are no phonemic palatalized velars in the core lexicon of Russian, while palatalized coronals and palatalized labials are a part of the underlying inventory. In addition, dorsal consonants are subject to the change of primary place of articulation (consonant mutation) when followed by front vowels in some morphologically derived forms, whereas coronals and labials in the same environments surface with secondary palatalization.

This composite place markedness scale by itself, however, does not provide any explanation for why assimilation happens in some clusters and not others. Some constraint has to trigger spreading. I suggest that spreading looks for better parsing, where secondary coronal must be dominated by the least marked primary place in a cluster. This suggestion is based on the assumption that consonant clusters are treated as complex place configurations where adjacent consonants behave as a single complex segment and secondary articulation may be associated, or parsed under either one of the adjacent primary places.³ To formalize the relations between secondary and primary articulation in a cluster I suggest a following constraint:

- (3) MAXIMIZE LICENSING: if a > b
 then f

² This hierarchy is similar to the Coronal Composite Unmarkedness scale suggested by Smolensky (1994) for the inventory analyses of complex places associated to a common root node.

³ In Zubritskaya (1994) I argue that the same composite place markedness hierarchy is responsible for the computation of all optimal place configurations in Russian and must be reformulated to allow for the notion of prosodic dominance and corresponding markedness: complex place dominated by a single root node in the case of complex segments, complex place dominated by a single syllabic node for onset/coda clusters, or a complex place dominated by a single PrWd node for heterosyllabic clusters. This distinction, however is not relevant for the present analysis, since the following discussion will concern only complex segments.

This constraint forces parsing of a secondary feature under the least marked of the two adjacent primary places. The feature does not migrate from one segment to another since faithful preservation of coindexing relations between a feature and its sponsoring root node is required.

At the outset of the sound change the assimilation was obligatory and the spreading constraint had to dominate the constraints that militate against complex place configurations parsed under the same root node. Note that the ranking between MAX and *DORS/COR' is irrelevant, since Dorsals were never targeted by assimilation. The immunity of velars to assimilation follows from the fact that Dorsal is the most marked place feature.

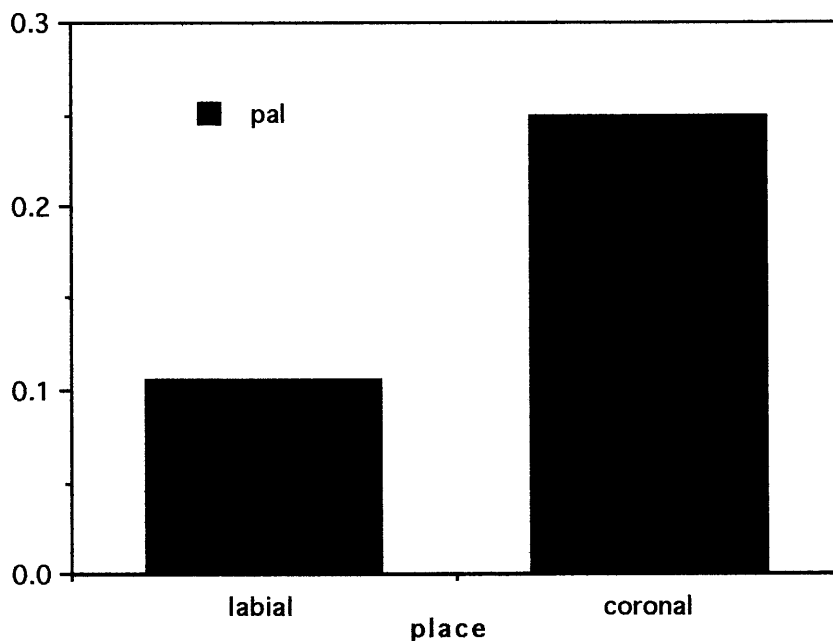
- (4) MAX
 *DORS » *LAB » *COR
 COR COR COR

As I have mentioned earlier this type of palatalization assimilation is disappearing in Russian. As most sound changes, however, it consistently shows a structured pattern of variation with different rates of assimilation loss in different environments. In the following discussion I will examine the frequencies of assimilation in different contexts and show that the directionality of palatalization loss is directly related to the markedness relations among functionally similar environments.

2.2. Place effects

The first striking characteristics of palatalization loss involves variable frequency of assimilation depending on the place of articulation of the target.

Graph 1. Rates of palatalization assimilation depending on the place of the target.



It is clear that assimilation is lost faster if the target is Labial. In OT this development finds a straightforward explanation if we adopt an assumption that a sound change is a restructuring of the constraint hierarchy, where the constraint that requires assimilatory spreading (MAX) is reranked below the whole family of constraints militating against complex place specifications:

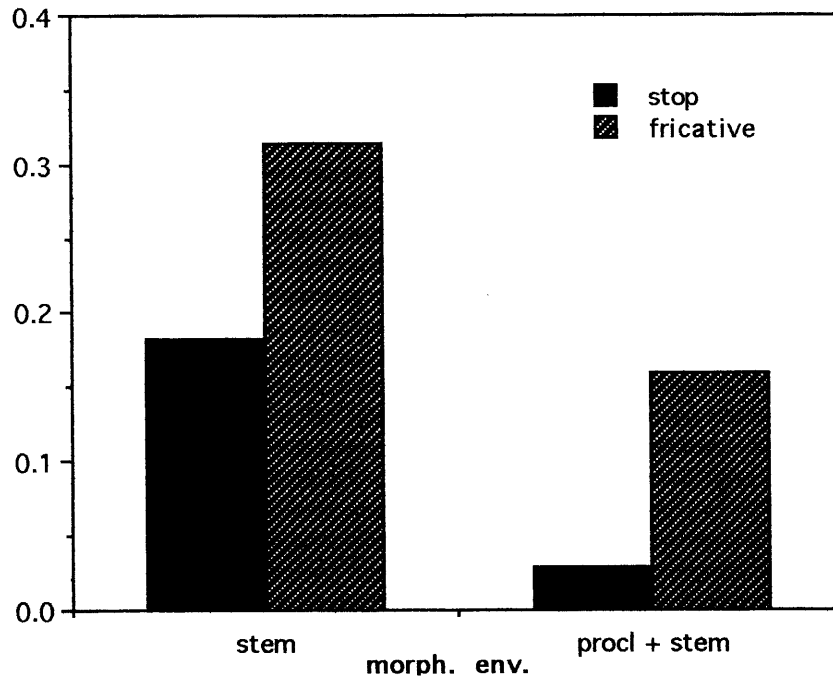
$$(5) \quad \begin{array}{c} \text{MAX} \\ \text{*DORS} \\ \text{COR} \end{array} \Rightarrow \begin{array}{c} \text{*LAB} \\ \text{COR} \end{array} \gg \begin{array}{c} \text{*COR} \\ \text{COR} \end{array}$$

The sound change in this model is the gradual weakening of a constraint that requires spreading and corresponding strengthening of a constraint family that bars complex segments. The gradience of sound change with respect to several segmental environments is directly predicted in the reranking model, since all environments are formally discrete and implicationaly ranked in the grammar. The rates with which one environment will favor or disfavor the newer forms is directly related to the markedness strength asymmetries between individual constraints within a constraint family. Note that in this model the restructuring of grammar is necessarily gradient, since any reranking of non-adjacent constraints (MAX and COR/COR' for example) must involve intermediate reranking (MAX and LAB/COR'). Correspondingly, if a loss of assimilation affected coronal segments then it must have also affected labial segments, but not vice versa. The implicational ranking of constraints in OT makes such directionality of the sound change not only natural but the only possible. No additional stipulations about the operation of the sound change are needed: the sound change is governed by the same regularities as other categorical processes of Russian phonology.

2.3. Sonority effects

The second important factor in the operation of this sound change involves the segmental characteristics of the target with respect to its relative sonority. Graph 2 shows the rates of assimilation depending on the sonority of the target in two morphological environments. The pattern is quite straightforward: fricatives retain palatalization better than stops. The observed effect can not be contributed to the structural similarity between interacting segments. Thus, for example, the rates of assimilation do not differ significantly in fricative + stop clusters vs. fricative + fricative clusters. Rates of assimilation loss solely depend on the relative sonority of the target.

Graph 2. Rates of palatalization assimilation depending on the sonority of the target.



Note that only [-sonorant] segments differing in continuacy values are considered. One troubling aspect of the assimilation loss with respect to the relative sonority of a target involves the exceptional behavior of liquids. Both /l/ and /r/ show no assimilation at all in the speech of younger informants, even though they are the most sonorant segments in the inventory. With /r/ an explanation lies, I believe, in the exceptional articulation of this segment. It is the only coronal segment of Russian articulated with a tip of the tongue and not with the front of the blade. If palatalization necessarily involves laminal articulation (raising of the front of the tongue blade) the articulation of /r/ is not directly compatible with secondary palatalization (Jones and Ward (1969)). With /l/ on the other hand, the explanation is found in the feature geometry of this segment. Laterals in Russian (and cross-linguistically) do not always behave like other coronal segments. Following the suggestion of Walsh (1994) I assume that laterals are doubly articulated segments (both coronal and dorsal) and may pattern with both Coronals and Dorsals. I would like to remind the reader that Dorsal consonants were not palatalized at the outset of the sound change. Hence, dorsal articulation of /l/ inhibits palatalization assimilation and the exceptional behavior of /l/ in palatalization loss is a direct consequence of its featural composition.

To account for the data in this graph, I suggest that secondary vocalic articulation is best parsed under more sonorant root nodes. I will formalize this suggestion as a number of constraints that prohibit secondary (vocalic) articulation under consonantal root nodes.⁴

- (6) SONORITY: *[-CONT] » * [+CONT]
- $$\begin{array}{c} \text{PL} \\ | \\ \text{COR} \end{array}$$

$$\begin{array}{c} \text{PL} \\ | \\ \text{COR} \end{array}$$

⁴ Mester and Ito (1994) have proposed a similar hierarchy in their analysis of Japanese palatal assimilation. They similarly suggest that secondary vocalic articulation is best parsed under the most sonorant root node.

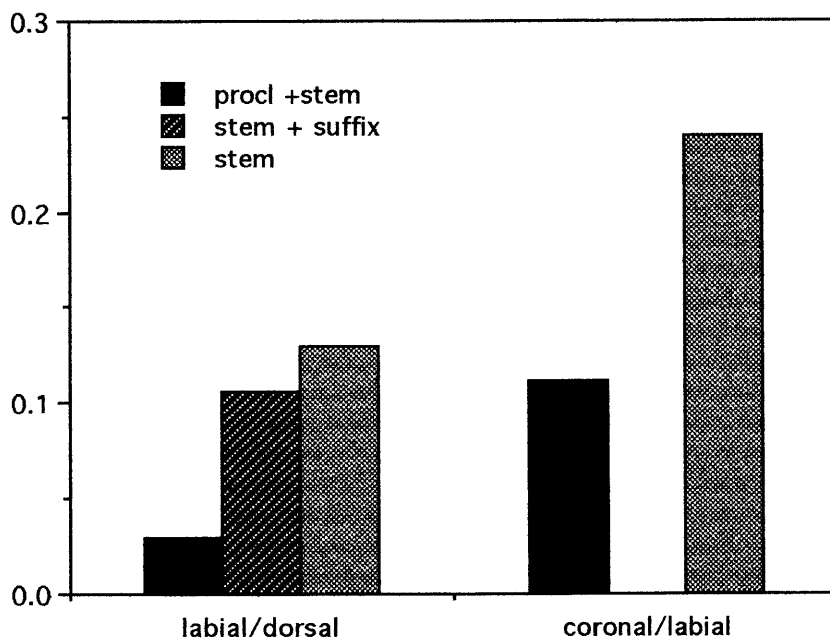
If these constraints are a part of Russian phonology, then palatalization assimilation loss must proceed as follows:

$$(7) \quad \text{MAX} \Rightarrow \begin{array}{c} *[-\text{CONT}] \\ | \\ \text{PL} \\ | \\ \text{COR} \end{array} \gg \begin{array}{c} *[\text{+CONT}] \\ | \\ \text{PL} \\ | \\ \text{COR} \end{array}$$

2.4. Morphological effects

Graphs 3 shows rates of palatalization assimilation in different morphological environments. Three environments are taken into consideration: stem-internal clusters, clusters at the boundary between a stem and a suffix, and clusters at the boundary between a proclitic (preposition/prefix) and a stem. If a graph is missing a particular morphological category, it means that this particular environment is not found in Russian. The major generalization emerges in a very straightforward manner: regardless of the overall rates of assimilation between different types of clusters, all of them show the same pattern of assimilation with respect to morphological environment. The highest rates of cluster assimilation obtain in monomorphemic forms, followed by stem + suffix boundary clusters, and the least assimilation happens between proclitics and stems. In terms of the dynamics of the sound change this means that assimilation is lost first of all at the boundary between proclitic and stem and only lastly in monomorphemic forms. The obvious candidate for the analysis of this directionality of the sound change in OT is Align constraint family (P&S 1993, M&P 1993). Languages are subject to a number of alignment constraints that require the edges of morphological and prosodic categories to coincide. These constraints are given below the graph.

Graph 3. Rates of palatalization assimilation for different morphological environments.



- (8) ALIGNMENT constraints (Prince & Smolensky 1993, McCarthy & Prince 1993ab):
 Align (MCat, PCat, Edge): the Edge of a morphological category MCat and the Edge of the prosodic category PCat must coincide.
 ALIGN L: Left edge of stem must coincide with the left edge of a Prosodic Word
 ALIGN R: Right edge of stem must coincide with the right edge of a syllable

Moreover, the ranking between Align L and Align R is fixed to reflect the cross-linguistic markedness asymmetry between the left and the right edges of a stem.

- (9) Align stem L, PrWd L » Align stem R, s R

The role of these constraints in Russian phonology is motivated quite independently from the facts of this particular sound change. Various categorical phenomena, such as cluster simplification, ‘jer’ vocalization, and vowel hiatus resolution, require an active role of ALIGNMENT constraints in Russian phonology.⁵ Why do the ALIGNMENT constraints come into play in the process of assimilation loss? Precisely because doubly linked structures incur violations of alignment. (In a strict reading, ALIGNMENT requires that all terminal (graph theoretic) nodes of a segmental form terminate in the same nodes of prosodic structure). The loss of palatalization assimilation can be characterized as a gradual strengthening of ALIGNMENT constraint family:

- (10) MAX ⇒
ALIGN L » ALIGN R

The universal markedness asymmetry between the ALIGNMENT constraints directly predicts the directionality of Russian sound change: since ALIGN L is ranked higher than ALIGN R, the assimilation is lost at higher rates at the left edge of a stem. Since no ALIGNMENT constraint is violated in monomorphemic forms, the loss of assimilation affects such forms the least. The implicational ranking of constraints makes a strong prediction for the reranking model of sound change: if assimilation is lost at the right edge of a stem then it is also lost at the left edge of a stem, but not vice versa.

2.5. Cluster simplification

The second sound change which shows the role of the same constraint family involves the incoming process of closure deletion in fricative + affricate clusters. This process is “fed” in a derivational terminology by a general process of coronal harmony:

- (11) r a z + ž e c → r a ž ž e č ‘to start a fire’
r a s + š i t' → r a š š i t' ‘to embroid’
r a s + č i s t i t' → r a š č i s t i t' ‘to clear up’

I assume that this type of assimilation occurs under the pressure of the Obligatory Contour Principle (McCarthy 1989), where the OCP violation is avoided by linking adjacent identical primary places. Assimilated forms have the following representation:

- (12) +cont (-cont) +cont
 <cor> cor
 f

⁵ The markedness asymmetry between the left and the right edges of the stem in Russian enforces non-incorporation of the proclitic material into a Prosodic Word with a stem and accounts for the exceptional resistance of the left edge of a stem to any kind of segmental or prosodic modification (Yearley (1994), Zubritskaya (1994)).

The doubly linked fricative + affricate clusters are undergoing a change and are replaced by a long post-alveolar fricative. This change involves a loss of the stop portion of an affricate and creation of a geminate structure⁶:

- (13) š č a s t'j e → š š a s t'j e 'happiness'
 p e r e n o s + č i k → p e r e n o š š i k 'carrier'
 s č e g o → š š e g o 'from what'

What triggers closure deletion? It is very likely that closure deletion is a consequence of articulatory overlap reinterpreted by younger speakers as an operation of a formerly latent constraint. This constraint can be formulated as a prohibition against linked structures with nonidentical stricture values.

- (14) ANTI-LINK: *G H where G ≠ H = stricture
 place

This constraint have been suggested by Selkirk (1991) as 'Homogeneous stricture linking'.⁷ One way of avoiding ill-formed linked structures is getting rid of an offending stricture feature and leaving [-continuant] unparsed. We must assume, however, that a faithfulness constraint PARSE was ranked above the ANTI-LINK in the Russian grammar at the outset of this sound change, since simplification of fricative + affricate clusters was prohibited in all environments including monomorphemic clusters. Let's suppose that during the sound change ANTI-LINK sneaks ahead of its superior and comes to dominate PARSE. This model of the sound change involves a simple reranking of two constraints:

- (15) Ranking at the outset of the sound change:
 PARSE » ANTI-LINK (no simplification is allowed)
- (16) Ranking after the sound change:
 ANTI-LINK » PARSE (simplification is allowed)

This sound change, however, can not be characterized simply as reranking of PARSE and ANTI-LINK. Like a loss of palatalization assimilation, this sound change shows gradience that must be analyzed as an operation of other constraints in Russian grammar.

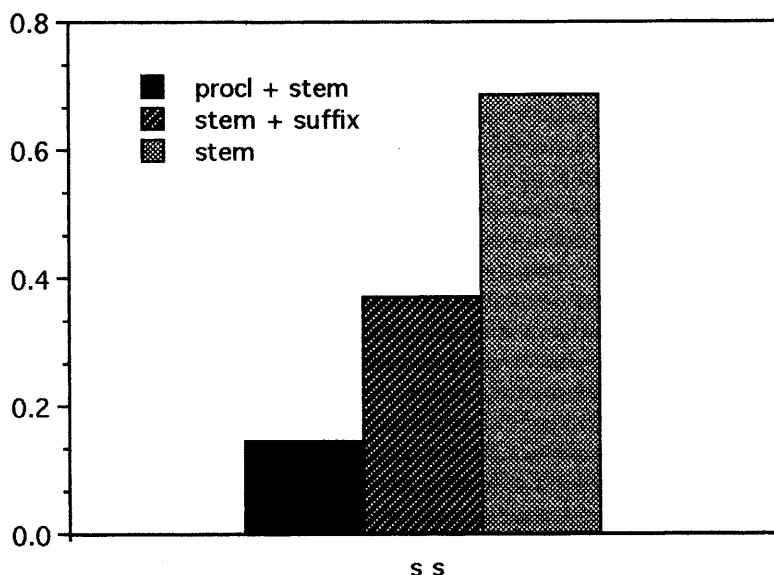
2.6. Morphological effects

The analysis of variable data shows that morphological environment consistently affects the operation of this sound change.

⁶ The fricative sequences formed as a result of closure deletion never undergo geminate simplification obligatory in some environments. The exceptional behavior of such surface geminates is, perhaps, related to the "visibility" of unparsed segmental material. If unparsed [-continuant] feature is still a part of representation, then two segments in such clusters are never identical.

⁷ Note that an alternative formulation of this constraint is possible based on Steriade's theory of closure and release: the prohibition against linked structures with three consecutive Aperture nodes may be one of the anti-structure constraints available in a grammar.

Graph 4. Rates of closure deletion depending on the morphological environment.



The most deletion occurs in monomorphemic forms, less across the stem + suffix boundary and the least across the left edge of a stem. This directionality of the sound change is readily explainable in OT. Since feature deletion is de-aligning (a feature lies at the edge of a morphological category and must be parsed just as the dominating root-node must be) it is sensitive to alignment constraints if these constraints are ranked higher in the hierarchy.⁸ (I assume that violation of ALIGN are cumulative (every instance of a linked or unparsed feature is an additional violation).

If the incoming sound change of closure deletion "overpowers" conflicting constraints and as a result gradually spreads in all environments, then the operation of this sound change can be represented as follows:

$$(17) \text{ ALIGN L} \gg \text{ ALIGN R, PARSE} \\ \leftarrow \text{ ANTI-LINK}$$

The implicational ranking of ALIGNMENT constraints makes the observed directionality of a sound change unavoidable. The incoming sound change must affect the least marked environment at a greater rate. In fact, the OT model predicts that the reverse situation where the incoming change shows higher rates of deletion at the left edge of a stem is impossible.

This sound change operates as a direct opposite of palatalization loss. While palatalization assimilation is in the process of disappearing, closure deletion is becoming "stronger". Regardless of the eventual outcomes of these sound changes, however, they both show the same conditioning environments. In a model of constraint reranking, loss of palatalization operates as a strengthening of ALIGNMENT constraints, while closure deletion

⁸ The ranking of ALIGN constraints above ANTI-LINK at the outset of this sound change is required independently if we are to preserve one ranking for all of the Russian grammar:

- a) ALIGN must dominate PARSE in Russian phonology for a coherent analysis of a number of categorical phonological phenomena.
- b) PARSE must dominate ANTI-LINK (at the outset of this sound change closure deletion was not allowed in any environment, including monomorphemic forms).
- c) If ALIGN \gg PARSE and PARSE \gg ANTI-LINK then ALIGN \gg ANTI-LINK.

operates as a weakening of ALIGNMENT constraints. The markedness ranking of individual constraints within an ALIGNMENT family in each case predicts the observed directionality of these sound changes. Moreover, in OT model of sound change it is the same markedness relations that are responsible for categorical phenomena of Russian phonology as well.

To summarize the discussion so far: The contention that whole families of functionally-related constraints are involved in the operation of the sound change makes possible a formal examination of gradualness and stratification of the sound change with respect to different linguistic environments. The same markedness relations that govern categorical phenomena determine the directionality of sound change in so far as they directly predict which environments will favor or disfavor the incoming newer forms. The speaker does not, in other words, have to learn the directionality of a sound change with respect to functionally similar environments since the operation of the sound change is determined by the universal markedness ranking. Modeling sound change in OT makes explicit predictions about the possible directionality of a sound change: if palatalization assimilation, for example, is lost in monomorphemic forms than it is also lost across prefix + stem boundary but not vice versa.

3. Modeling variation in OT

I would like to turn now to the second set of issues raised by variable data. The examination of palatalization loss has shown that several constraint families may be simultaneously involved in the process of sound change. If sound change in OT must be modeled as constraint reranking, then these non-related constraint families must be reranked simultaneously. In OT model of strict domination and categorical ranking, this essentially requires adjacency of conflicting constraints in the hierarchy. Moreover, the simultaneous access of several constraints to the operation of a sound change in the model of reranking is possible only if these constraint families occupy the same place in the hierarchy. For example, at the outset of palatalization loss, the ALIGNMENT family and the *COMPLEX PLACE constraint family, which are not related in any formal sense, must both be situated below the spreading constraint MAX so that the latter may enforce assimilatory linking in all environments. More importantly, these two families must occupy the same place in the overall constraint hierarchy, so that reranking of MAX means reranking with respect to both of them. This is indeed the case in Russian palatalization loss. Crucially, there is no categorical process in Russian grammar that requires ranking between ALIGNMENT and *COMPLEX PLACE. (They never impose conflicting requirements and no segmental process in Russian requires such ranking even indirectly). This distribution of constraints in Russian grammar allows a partial explanation for the simultaneous involvement of overlapping constraint families in the process of the sound change. The process of reranking is represented schematically:

- (18) MAX \Rightarrow
- | | | |
|-----------------|---|-----------------|
| *LAB/COR | » | *COR/COR |
| ALIGN L | » | ALIGN R |
| *[-CONT]/PL/COR | » | *[+CONT]/PL/COR |

Note that no extension of OT formalisms was necessary for this part of the analysis; in fact, non-ranking of constraints is very much a part of many OT models, where no categorical phonological process can determine the ranking. Such a notion of non-ranking or "tie" between several non conflicting constraint families, however, only explains why several constraint families can influence the operation of a sound change simultaneously, providing no explanation of data gradience and variability. As I mentioned in the beginning of the

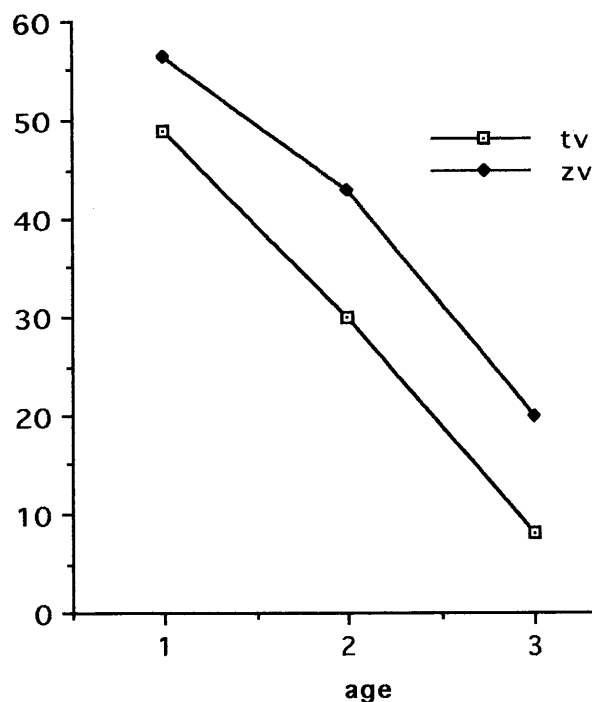
paper, OT does not differ from other standard models of phonological representation in the assumption that phonological patterns are essentially nonoptional. Since the constraint ranking is categorical a given input can have only one fixed output in an OT grammar. Optionality in the choice of outputs for the same input has to be modeled in OT as a competition among grammars. While the contention that sound change must ultimately involve two different rankings is not problematic, the status of variation as grammar competition is. The preservation of categorical ranking entails a step-by step reranking of an individual constraint below each of the constraints in a family so that, for example, we can have the following three rankings for two possible environments:

- (19) MAX » *LAB/COR » *COR/COR
 *LAB/COR » MAX » *COR/COR
 *LAB/COR » *COR/COR » MAX

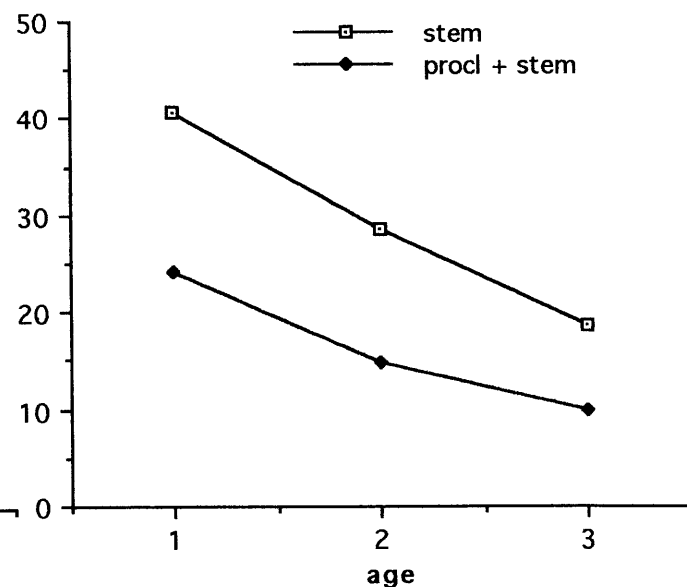
Such model of language change would entail that the gradient effects of individual constraints within constraint families during the period of sound change are the result of grammar competition. Correspondingly, the probabilities of data distribution must depend on probability distribution among these competing grammars. What implications would such a model have for the variable data discussed in this paper? For the single process of palatalization assimilation loss the number of competing grammars is unrealistically large once we calculate that, for each segmental or morphological environment, there is a binary option of ranking (assimilation versus nonassimilation). There are n environments, and there are 2^n grammars. There are 6 possible environments for the palatalization data: hence we need 2^6 grammars. This is not a good result since the number of grammars will increase even further as soon as we examine the operation of this sound change in other environments. I have mentioned that palatalization assimilation loss also affects homorganic clusters in Russian, which bring in additional factors and additional environments. The total of possible environments for palatalization loss in all cases is about 18 (Zubritskaya (1994). The average Russian speaker must be capable of choosing among an astronomical number of competing grammars for a single sound change. Note that in a model of variation as grammar competition we must assume that all possible grammars are available throughout the duration of a sound change. A step-by-step reranking model which predicts that a sound change starts in the most marked environment and affect a less marked environment only after it is completed in the most marked environment simply can not model variation unless all relevant grammars are competing for the most of the duration of a sound change. In fact sound change never affects environments one after another as the gradual reranking model will predict. Consider, for example, graphs 5 and 6 which show the rates of palatalization assimilation loss for three age groups. Each graph demonstrates the loss of palatalization assimilation in two words with contrasting phonological or morphological environments. In graph 5 the differences between the assimilation of the target stop versus target fricative in two monomorphemic forms are examined. Graph 6 illustrates the loss of assimilation for the same cluster in two morphological environments.

Graph 5. Fricative vs. stop

(20) T'v'er' - name of the city
s'+ p'iset - 'write down'

**Graph 6.** Stem vs. proclitic + stem

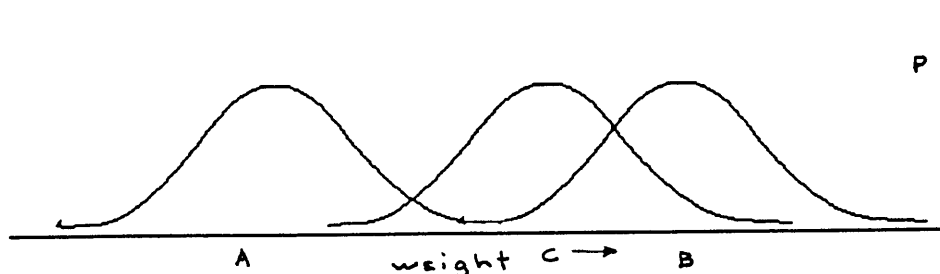
(21) s'p'i - 'sleep'
z'v'er' - beast



These graphs clearly show that the sound change affects two relevant environments simultaneously at all points in time (so that the rates of palatalization are decreased in tandem for both environments). This mechanism of linguistic change was discussed by Kroch as a Constant Rate Effect (1989), who showed that in syntactic change, a single optional rule is introduced simultaneously in several environments and changes its frequency of use in all these environments at the same rate, with the rate to which one environment favors or disfavors newer forms constant through time. These graphs show that similarly, the rates of palatalization decrease in tandem for all environments and the degree to which each particular environment favors newer forms is the same at every point in time. While these figures show exactly the type of distribution discussed by Kroch, the Constant Rate Effect has not been proved to be a fundamental characteristics of phonological as well as syntactic change. At this point, it requires further study, beyond the scope of this paper. In this paper I simply wanted to show that phonological change does affect all relevant environments simultaneously for most of its duration. Crucially, to model variation as reranking, we have to assume that all possible rankings are available as competing grammars for all speakers and the range of possible grammars is known to the speaker. Moreover, even given the assumption that all possible competing grammars are available throughout the duration of a sound change, a model of grammar competition does not provide any explanations for the observed frequencies of data distribution, since any frequencies at all can be in principle modeled given several grammars. The actual observed frequencies are not predictable and are not in any way related to the structure of grammar. To explain the Constant Rate Effect, for example, we have to stipulate that the rates of grammar mixture stay constant for all points in time, and an increase in a choice of a particular ranking for one environment leads to a proportional increase in a choice of a similar ranking for another environment. It is quite clear that the model of grammar

competition can neither predict these characteristics of sound change nor explain the observed frequencies of data distribution.

We are faced with a seemingly unavoidable situation: a sound change clearly involves a state of competition of a single constraint with whole constraint families. To explain and predict the directionality of the sound change with respect to markedness and naturalness and preserve the categorical ranking, all of the individual constraints within each constraint family must be directly incorporated into the model of reranking. This incorporation, however, results in competing grammar model which does not provide any explanations for the observed frequencies of data distribution and requires an unrealistically large number of grammars. Moreover, Mark Liberman (1994) notes that in a model of grammar competition, the same phenomena of linguistic markedness is modeled in categorical processes as an effect of particular constraints, but in variable processes of sound change by a totally different mechanism of grammar weighting. The solution lies, perhaps, in the examination of sound change as (Figure 1) a gradual movement of a single distribution (constraint C) along a weight scale, where it competes with two (or more) functionally related constraints (A and B), whose weight remains constant.



If the probability of the appearance of a particular form can be represented as a normal distribution curve, then for the categorical phenomena, constraint weight would ensure that the curves overlap is so insignificant that the ranking is categorical. If the sound change is a gradual movement of a particular constraint (distribution) in the hierarchy, then a partial overlap of several competing constraints is possible as in Figure 1, where all curves have the same means of deviation. The probability of occurrence of one form versus the other can be directly modeled as a difference in weight of competing constraints. This model of language change essentially postulates only one grammar for every point in time, where for every such point in time the frequencies of data distribution are directly predictable. This schematic representation of the sound change can be directly translated into the Graphs 5 and 6, where the differences in rates to which two environments favor the newer forms are directly related to the fixed differences in weight between the functionally similar constraints (here constraints A and B). Hence these rates must stay constant through time (i.e. Constant Rate Effect). The differences in the overall rates of newer forms (loss of palatalization, for example), on the other hand are directly determined by the degree of distributions overlap (C with A and B). Note that this while this model introduces the notion of constraint weight it does not mean that several lower ranked constraints can somehow outweigh higher ranked ones, it only allows the probability of data distribution during the period of the sound change (only the period of changing weight), to be directly determined from the degree of distributions overlap. While this suggestion is tentative and requires further elaboration, it allows a certain relaxation and modification of current OT formalisms. Such modifications, I believe, are necessary if the analyses of sound change and other variable phenomena are to be incorporated into OT model of phonology.

References

- Bailey, C.-J., 1972. The Integration of Linguistic Theory: Internal Reconstruction and the Comparative Method in Descriptive Analysis. in Stockwell and Macaulay, eds., *Linguistic Change and Generative Theory*, Indiana Univ. Press, 22-31.
- Bailey, C.-J., 1973. *Variation and Linguistic Theory*. Arlington, Va.: Center for Applied Linguistics.
- Ganiev, Z. V., 1971. On Pronunciation of Workers Born in Moscow. in *Razvitie fonetiki sovremennogo russkogo jazyka*, 2. Visotskij, Panov, Sidorov eds, Nauka., Moscow, Russian Language Institute. Academy of Sciences, USSR.
- Hume, E. (1992). *Front Vowels, Coronal Consonants, and their Interaction in Nonlinear Phonology*. Doctoral Dissertation. Cornell University, Ithaca.
- Kroch. A. S., 1989a. Function and Grammar in the History of English Periphrastic *do* in Fasold R. and Schiffrin D., (eds.), *LVC. (Current Issues in Linguistic Theory 52)*. Philadelphia: John Benjamins. 133-172.
- Krisin. ed. 1974. *Russkij Jazyk po Dannym Massovogo Obsledovanija*. Nauka, Moscow, Russian Language Institute. Academy Of Sciences. USSR.
- Labov, W. 1994. *Principles of Linguistic Change. Volume 1. Internal Factors*. Part A, E. Blackwell, Oxford UK and Cambridge USA.
- Lieberman, M. 1994. *Optimality and Optionality*. ms. University of Pennsylvania.
- McCarthy, J. & Prince, A. 1993. *Prosodic Morphology 1: Constraint Interaction and Satisfaction*. Manuscript, U. of Massachusetts, Amherst, and Rutgers University.
- Prince, A. & Smolensky P. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. Manuscript, Rutgers U. and the University of Colorado at Boulder.
- Smolensky P. 1993. *Harmony, Markedness and Phonological Activity*. Handout, Rutgers Optimality Workshop.
- Visotskij, Panov, Sidorov eds. 1966. *Razvitie Fonetiki Sovremennogo Russkogo Jazyka*. 1. Nauka., Moscow, Russian Language Institute. Academy of Sciences, USSR.
- Zubritskaya, K., 1994. *Cluster Simplification and Alignment in Russian*, University of Pennsylvania, ms.

Linguistics Department
 619 Williams Hall
 U. of Pennsylvania
 36th Spruce St.
 Philadelphia, Pa 19104