ABSTRACT

Title of dissertation: THE PHONOTACTICS AND PHONOLOGY OF OBSTRUENT CLUSTERS IN OPTIMALITY THEORY

Frida Morelli, Doctor of Philosophy, 1999

Dissertation directed by: Professor Linda Lombardi
Department of Linguistics

In this dissertation, I present a typology of obstruent clusters and argue that the systematic patterns of occurrence of these clusters must be explained in sonority-independent terms. I argue that there are two dimensions along which generalizations can be made; one is the dimension where the feature [continuant] is relevant and the other is the place dimension.

On the continuancy dimension, I claim that markedness relationships exist among the four types of obstruent clusters, i.e. fricative-stop, fricative-fricative, stop-fricative and stop-stop. In particular, I argue that fricative-stop clusters are the unmarked type of obstruent clusters. Their unmarkedness is assessed against a system of constraints on segment sequencing on the basis of a strategy of analysis that derives universal markedness relationships without fixed rankings. I call this
strategy the *Subset Strategy*. Modern Greek and Nisqually provide evidence for my proposal, as does a cross-linguistic survey of obstruent cluster patterns.

On the place dimension, I show that the typology observed can best be understood via a system of constraints that favors faithfulness to place in release positions. Relevant data from English, German, Dakota and Takelma are presented and analyzed.

Finally, I argue that, s+STOP clusters are the best-formed of all the obstruent clusters because they are unmarked along both dimensions. I show that the special phonological behavior often associated with these clusters follows from the fact that they are unmarked within the dimension of obstruent clusters, and not because they are marked within the dimension of core clusters.
THE PHONOTACTICS AND PHONOLOGY OF OBSTRUENT CLUSTERS IN
OPTIMALITY THEORY

by

Frida Morelli

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland at College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
1999

Advisory Committee:
Professor Linda Lombardi, Chair/Advisor
Professor Laura Benua
Professor Luigi Burzio
Professor Mary Ellen Scullen
Professor Paul Smolensky
© Copyright by

Frida Morelli

1999
DEDICATION

To my mother, Mirella, and my late father, Antonio.
ACKNOWLEDGMENTS

It is particularly hard to find words to express my gratitude to my advisor, Linda Lombardi, for her academic as well as personal support towards the completion of this dissertation. She has not only provided me with invaluable feedback on my work but also with constant encouragement during difficult times.

I am also indebted to the rest of my committee, Laura Benua, Luigi Burzio, Mary Ellen Scullen, and Paul Smolensky for their comments, advice, and feedback. Special thanks to Laura Benua for her input on my work, for the Phonology Reading Group at her house, and nice conversations on topics other than phonology. I would also like to thank Paul Smolensky for helping me understand the intricacies of Optimality Theory during his seminars at Johns Hopkins and for those appointments that I was lucky enough to steal from his busy schedule. I especially thank him for sticking with me until the end and for his kind words of encouragement during rough times.

Thanks are also due to the rest of the faculty of the Department of Linguistics at the University of Maryland and Kathi Faulkingham for their support during my doctoral study, as well as to Dee Ann Holinsky, Charlie Jones and Steven Weinberger at George Mason University for introducing me to the study of Linguistics.

Thanks also to John McCarthy, Donca Steriade and audiences at HOT 97, LASSO 97, NELS 28, LSA 98, “The Syllable: Typology and Theory” in
Tübingen (1998), the “NSF Workshop on Syllable Structure and Gesture Timing” at Ohio State University (1998), RUMD 97 and 99, for comments and feedback during the development of this dissertation. I would also like to thank all the linguists who have replied to a query that I posted on the Linguist List, and especially Ellen Kaisse for providing crucial data for the analysis presented here.

Next, I would like to thank my fellow funologists at the University of Maryland, Haruka Fukazawa, Pat Hironymous, Viola Miglio, Bruce Morén and Caro Struijke, for sharing all the ups and downs of our doctoral study and much more. Thanks to Mits Ota for joining us from Georgetown and sharing these moments with us too.

For their support, encouragement and friendship I would like to thank Stefania Amodeo, Andrea Gualmini, Shirley Mani, Luisa Meroni, Mitsue Motomura, Acrisio Pires, Riccardo Sala, Elizabeth Stranges, Eden Teklu and Eileen Timothy.

Special thanks go to Mark Yancey for providing me with emotional support and for rescuing my work from the many computer crashes that have haunted my dissertation. I would also like to thank Daniele Struppa for giving me the opportunity to start my career as a linguist in the United States.

Least, but not last, I would like to thank my mother, Mirella, and my sister, Daniela, for everything they have done to allow me to make this dissertation come to an end!
TABLE OF CONTENTS

CHAPTER 1: CLUSTER PHONOTACTICS AND THE SONORITY SEQUENCING PRINCIPLE

1.1 Introduction .................................................................................................................. 1

1.2 Sonority and the Sonority Sequencing Principle ......................................................... 4
  1.2.1 Sonority .................................................................................................................. 4
  1.2.2 Sonority Scales ..................................................................................................... 5
  1.2.3 The Sonority Sequencing Principle ....................................................................... 7

1.3 Optimality Theory ........................................................................................................ 9
  1.3.1 Basic Architecture ................................................................................................. 9
  1.3.2 Markedness as Harmony ...................................................................................... 10
  1.3.3 Implicational Universals ....................................................................................... 16

1.4 Background Assumptions ............................................................................................ 20

1.5 The Sonority Sequencing Principle and Optimality Theory ...................................... 22

1.6 Summary of the chapter. ............................................................................................. 28

CHAPTER 2: THE ONSET TYPOLOGY

2.1 Introduction .................................................................................................................. 30

2.2 Methodology ................................................................................................................. 32
  2.2.1 Obstruent Clusters: Definition ............................................................................. 32
  2.2.2 The Status of Affricates ...................................................................................... 33
  2.2.3 SF Sequences: Clusters or Singletons? ................................................................. 35
  2.2.4 The Sampled Languages ..................................................................................... 36
  2.2.5 Syllable Onsets .................................................................................................... 38
  2.2.6 Morphologically Complex Clusters ...................................................................... 39
  2.2.7 Non-native Phonotactics ..................................................................................... 39

2.3 Onset Generalizations .................................................................................................. 41

2.4 Sonority and the Typology .......................................................................................... 45
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHAPTER 4: PLACE GENERALIZATIONS</td>
<td>118</td>
</tr>
<tr>
<td>4.1 Introduction</td>
<td>118</td>
</tr>
<tr>
<td>4.2 Onset Place Restrictions</td>
<td>119</td>
</tr>
<tr>
<td>4.3 Release-sensitive Faithfulness and Place Neutralization</td>
<td>120</td>
</tr>
<tr>
<td>4.4 Case Study III: English</td>
<td>124</td>
</tr>
<tr>
<td>4.4.1 The English Obstruent System</td>
<td>125</td>
</tr>
<tr>
<td>4.5 Case Study IV: German</td>
<td>130</td>
</tr>
<tr>
<td>4.5.1 The Obstruent System</td>
<td>130</td>
</tr>
<tr>
<td>4.5.2 Phonotactics of Onset Obstruent Clusters</td>
<td>131</td>
</tr>
<tr>
<td>4.5.3 Analysis of Onset Obstruent Clusters</td>
<td>134</td>
</tr>
<tr>
<td>4.6 Case Study V: Onset Place Asymmetries in Delaware</td>
<td>138</td>
</tr>
<tr>
<td>4.7 Case Study VI: Takelma</td>
<td>142</td>
</tr>
<tr>
<td>4.7.1 The facts about Takelma</td>
<td>143</td>
</tr>
<tr>
<td>4.7.1.1 An Additional Fact about Takelma</td>
<td>146</td>
</tr>
<tr>
<td>4.8 Licensing by Cue</td>
<td>146</td>
</tr>
<tr>
<td>CHAPTER 5: MEDIAL SYLLABIFICATION OF s+STOP ONSETS</td>
<td>154</td>
</tr>
<tr>
<td>5.1 Introduction</td>
<td>154</td>
</tr>
<tr>
<td>5.2 s+STOP Onsets and the Sonority Sequencing Principle</td>
<td>155</td>
</tr>
<tr>
<td>5.3 s+STOP Onsets: the Unmarked Phonotactics</td>
<td>157</td>
</tr>
<tr>
<td>5.4 Case Study VII: Standard Italian</td>
<td>161</td>
</tr>
<tr>
<td>5.4.1 Cluster Phonotactics</td>
<td>162</td>
</tr>
<tr>
<td>5.4.2 Syllabification</td>
<td>165</td>
</tr>
<tr>
<td>5.4.3 Morphological Alternations</td>
<td>174</td>
</tr>
<tr>
<td>5.5 Sanskrit Perfect Reduplication</td>
<td>183</td>
</tr>
<tr>
<td>5.5.1 The Analysis of Reduplication of Obstruent Clusters</td>
<td>185</td>
</tr>
<tr>
<td>CONCLUSION</td>
<td>193</td>
</tr>
<tr>
<td>APPENDIX I</td>
<td>197</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>224</td>
</tr>
</tbody>
</table>
CHAPTER 1

CLUSTER PHONOTACTICS AND THE SONORITY SEQUENCING PRINCIPLE

1.1 Introduction

Languages of the world differ in their syllable phonotactics. Some languages are extremely restrictive and only allow CV sequences; others allow more complex structures both in the margins and nuclei. Across languages, segments are organized into well-formed sequences according to universal principles of segment sequencing. The organization of segments within the syllable, and across syllables, is traditionally assumed to be driven by principles of sonority, a property that ranks segments along a hierarchy from most sonorous to least sonorous. A number of strong cross-linguistic tendencies on the distribution and sequencing of segments is explained with reference to the Sonority Hierarchy. Principles such as the Sonority Sequencing Principle, introduced as early as the 19th century by Sievers (1881), and later by Jespersen (1904), explains, for instance, the tendency, within a syllable, of more sonorous segments to stand closer to the syllable peak than less sonorous ones. The Minimum Sonority Distance Principle, introduced by Harris (1983), explains language-specific patterns of consonant clustering by proposing that segments combine on the basis of their relative distance on the sonority scale. Sonority-based principles are not limited to intrasyllabic sequences, the Syllable Contact Law (Murray and
Vennemann, 1983), applies to intersyllabic segment sequences. It holds that the preferred contact between two adjacent syllables is when the segment ending the first syllable is higher in sonority than the segment beginning the second syllable.

Although sonority-based principles of segment organization capture the most common patterns found across languages, they, however, are not without exceptions. Clusters that are not predicted by sonority-based generalizations are relatively frequent across languages. As an example, initial s+stop clusters are commonly found across a number of unrelated languages, despite the fact that they constitute violations of the Sonority Sequencing Principle. The main tendency in the phonological tradition has been to account for the occurrence of such violations by means of special syllabification rules or representations, which would make these sequences immune to sonority. The attempt to reconcile the occurrence of sonority violations with sonority-based principles was the main motivation behind these approaches. In this dissertation, I argue that, at least in the case of obstruent clusters, there is no need to stipulate any special rule or representation to justify their immunity to the principles of sonority. I defend the hypothesis that obstruent clusters are not constrained by sonority principles. I show that the generalizations observed can only be explained in sonority-independent terms, which I formalize under Optimality Theory (Prince and Smolensky 1993). Moreover, I argue against the view that the asymmetric behavior shown by certain obstruent clusters in a number of languages is evidence for the special status of such clusters. On the contrary, I show that their different
behavior results from the interaction of basic syllable structure constraints as well as independent markedness considerations.

This dissertation has two main goals. One goal is purely empirical and its main purpose is to contribute to the understanding of universal principles of syllable phonotactics and segment patterning. A number of facts about the distribution and co-occurrence restrictions of obstruent clusters that result from a cross-linguistic study are presented and analyzed. The other goal is instead theoretical. From this point of view, the dissertation contributes to the understanding and implementation of the tools available in Optimality Theory by providing an explicit formalization of a technique of analysis, referred to as The Subset Strategy. This technique will be used repeatedly to capture the markedness relationships that I argue exist among the different types of obstruent clusters, as well as the implicational universals that follow from the systems of constraints. The dissertation also provides a detailed discussion and implementation of the Harmonic Bounding Argument, and addresses other current theoretical issues within Optimality Theory. In particular, I discuss the property of Strong Harmonic Completeness. I show how different dimensions of markedness can give rise to harmonically incomplete languages in a typology that is itself strongly harmonically complete. Finally, the dissertation contributes to the understanding of how phonetics and phonology may interact in the characterization of phonological grammars by showing how a number of
constraints which explicitly refer to phonetic facts can contribute to the understanding of phonotactic patterns.

In the rest of this chapter, I will briefly review some of the literature on sonority and sonority-based accounts of segment clustering. I will consider whether certain conclusions reached in previous literature still hold in view of the current framework of Optimality Theory. I will, moreover, provide arguments for why a satisfactory explanation of the well-formedness of the clusters that violate sonority cannot be formulated in sonority-dependent terms.

1.2 Sonority and the Sonority Sequencing Principle

1.2.1 Sonority

Although the notion that segments are ranked along a scale on the basis of their sonority is broadly accepted, the question of what sonority is and how it could be defined still remains a highly controversial issue, both in the phonetic and in the phonological literature. From a phonetic point of view, researchers disagree on whether a single phonetic parameter should be used to define sonority, i.e. perceptual salience or loudness of a particular sound (Ladefoged 1982, 1993); or the amount of airflow in the resonance chamber (Bloch and Trager 1942, Goldsmith 1995); or whether it should be interpreted in terms of multiple phonetic parameters (Ohala and Kawasaki 1984; Ohala 1990; Butt 1992). In the phonological literature the issue revolves, instead, upon whether sonority should be a phonological primitive in the form of a multi-valued feature (Foley 1972;
Hankamer and Aissen 1974; Selkirk 1984), or whether it should be derivable from the more basic binary features of phonological theory (Clements 1990). Another strategy, instead, is not to deal with the nature of sonority itself, but rather derive the relative sonority of each segment on the basis of their occurrence within a syllable. In other words scales are constructed on the basis of the observed patterns of syllable organization in a language specific way (Steriade 1982; Davis 1990).

1.2.2 Sonority Scales

The many different approaches proposed to derive sonority have led to the proposal of a number of competing scales in the literature. The main issue is whether sonority scales are universal, i.e. there is only a single universal scale common to all languages (Selkirk 1984; Clements 1990; Butt 1992); or whether sonority scales are language-specific and languages have a certain degree of freedom in the assignment of sonority values to their segments (Steriade 1982). Sonority scales with fixed universal values mostly refer to the major natural classes of sounds. Finer distinctions among segments are derived by means of sonority-independent parameters, i.e. voicing, coronality etc.. Clements' universal sonority scale, for example, for nonsyllabic segments only consists of the four major natural classes of sounds (obstruents, nasals, liquids and glides) ranked from least sonorous to most sonorous, as in (1) below:

\[(1) \quad O < N < L < G\]
Butt's sonority scale differs slightly from Clements’ in that he assigns a different value to voiceless and voiced obstruents. His universal sonority scale consists of the following ranking:

(2) Voiceless O < Voiced O < N < L < G < V.

Selkirk (1984) assumes even further distinctions among the obstruents and the liquids and proposes the following universal sonority scale for non-syllabic segments:

(3) p, t, k < b, d, g < f, θ < v, z, ð < s < m, n < l < r

As noted by Steriade (1982), the problem with Selkirk's proposal is that different languages seem to assign contradictory values to the same entries on the scale. Steriade proposes, instead, that languages enjoy a certain level of freedom in the assignment of sonority values to their segments. Clements argues, however, that allowing the sonority scale to vary across languages seriously undermines its explanatory power. Clements writes: “… increasing the number of ways in which the sonority hierarchy can accommodate potential exceptions, will reduce the number of cross-linguistic generalizations that it accounts for”. As a matter of fact, both Clements and Butt argue that most of the apparent evidence for language particular variation in the sonority scale comes from observations that can be explained in ways which are sonority independent and should not count in the formulation of the scale (Clements 1990; Butt 1992).

For the purpose of this study, it is crucial that fricatives and stops constitute a single class with respect to sonority. In particular, in Section 2.3, I
argue that splitting the obstruents into separate classes with respect to sonority not only does not solve the problem, but actually makes the wrong typological predictions. I moreover argue that the generalizations for obstruent clusters and their patterning across languages must be stated in sonority-independent terms. Sonority not only does not explain the facts observed but makes also the wrong predictions.

1.2.3 The Sonority Sequencing Principle

Despite the lack of agreement on the nature of sonority itself, and the way sonority scales are constructed, its role in deriving some of the most common restrictions on segment sequencing is uncontroversial. One of the most general cross-linguistic patterns of syllable phonotactics is the generalization that in any syllable the segment ranking highest on the sonority scale constitutes the peak of the syllable. All the other segments are organized around the nucleus in such a way that the more sonorous segments are closer to the peak and the less sonorous ones are further away from it. This generalization, known in the literature as the Sonority Sequencing Principle (henceforth SSP), was noticed early on by Sievers (1881), Jespersen (1904), Sausurre (1914) and Grammont (1933). More recently, researchers such as Hooper (1976), Kiparsky (1979), Steriade (1982), Selkirk (1982), Clements (1990) have attempted to provide formal characterizations of the SSP.
Although the validity of the SSP in phonological theory is uncontroversial, the existence of clusters that do not conform to the pattern prescribed by the principle undermines its universality within theories in which constraints are not violable. As a matter of fact, given the occurrence of clusters that do not conform to this generalization, the SSP is best characterized as a universal tendency rather than an absolute universal. Within the framework of lexical phonology (Pesetsky 1979; Kiparsky 1982; Mohanan 1982), in which different levels of representations are allowed, the question arises at what linguistic level the SSP holds. Steriade (1982) and Clements (1990), for example, argue that the SSP only holds at the level of core syllabification, i.e. the level where the cyclic or lexical syllabification rules apply. Post-cyclic syllabification rules, in their proposal, are not constrained by relative sonority. More complex clusters are created by later adjunction rules applying at the periphery of the syllabification domain. By restricting the domain of application of the SSP to the level of core syllabification, both authors attempt, in one way or another, to preserve the universality of the SSP at the level where the principle applies.

Within Optimality Theory, the issue of the universality of the SSP does not arise due to the architecture of the framework, as I will show in Section 1.3. Before turning to such a discussion, I will introduce the basic architecture of Optimality Theory with particular attention to some evaluation procedures that will be used in the rest of the dissertation. In particular, I will focus on markedness within OT and introduce a strategy of analysis, that I call the Subset
Strategy. This strategy allows to derive markedness relations in the case in which no universal rankings can be determined. Moreover, in Section 1.4 I will provide a discussion of the Sonority Sequencing Principle and its role within a theory such as OT. This section does not contain the core of the proposal, but is rather intended to provide the background for the idea that obstruent clusters are different from other types of clusters and represent a unique phenomenon.

1.3 Optimality Theory

1.3.1 Basic Architecture

Optimality Theory (Prince and Smolensky 1993) is a theory of violable, universal constraints and their interaction. The basic architecture of OT can be represented in the following diagram adapted from Smolensky (1995):

(4) Architecture of Optimality Theory
Given an input, Gen produces an infinite number of possible output candidates, which are evaluated for harmony against Con, i.e. the set of violable universal constraints ranked on a language particular basis. The candidate that best satisfies the constraint system of the language in question is selected as the optimal surface form by H-eval. H-eval is a function that evaluates the violations incurred by the candidates and selects the most harmonic candidate on the basis of constraint violations. The most harmonic candidate corresponds to the candidate that best satisfies the constraint hierarchy. Best satisfaction of a constraint system is determined on the basis of satisfaction of higher ranked constraints at the expenses of violations of lower ranked constraints.

1.3.2 Markedness as Harmony

Built within the basic architecture of OT is a formal theory of markedness. Markedness corresponds to dis-harmony, and markedness relations among forms are expressed in terms of harmonic orderings of forms. Determining harmonic orderings consists in establishing the relative harmony of each output candidate on the basis of its constraint violations. In the case of single binary constraints, harmony evaluation of candidates consists in determining whether a candidate violates a given constraint or not. Since marks are by definition anti-harmonic, it follows that a candidate $\alpha$ that satisfies a constraint C is more harmonic (or unmarked with respect to C), than a candidate $\beta$ that violates it (which is therefore marked with respect to C). The harmonic ordering, or markedness
relationship between candidates $\alpha$ and $\beta$ is expressed as $\alpha \succ \beta$, where "$\succ$" means "more harmonic than". Harmonic ordering evaluation by a single binary constraint is shown in the tableau below:

(5) $\alpha \succ \beta$ with respect to $C$

<table>
<thead>
<tr>
<th></th>
<th>$C$</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. $\alpha$</td>
<td></td>
</tr>
<tr>
<td>b. $\beta$</td>
<td>*</td>
</tr>
</tbody>
</table>

Prince and Smolensky (1993) show that more complex cases of candidate evaluation by single non-binary constraints or entire constraint systems in domination hierarchies are reducible to the simple case of single binary constraints. This is because of the evaluation strategy that cancels common marks and evaluates candidates on the basis of the unshared marks. This strategy is called the $\textit{Cancellation Lemma}$ and is defined in Prince and Smolensky (1993) as follows:

(6) $\textit{Cancellation Lemma}$

Suppose two structures $S_1$ and $S_2$ both incur the same mark $^*m$. Then to determine whether $S_1 \succ S_2$, we can omit $^*m$ from the list of marks of both $S_1$ and $S_2$ (‘cancel the common mark’) and compare $S_1$ and $S_2$ solely on the basis of the remaining marks. Applied iteratively, this means we can cancel all common marks and assess $S_1$ and $S_2$ by comparing only their unshared marks.
The procedure is illustrated in tableau (7), where C’ stands for a gradient constraint, i.e. a constraint that assigns multiple violations to candidates. Shared violations are included in square brackets:

(7) \( \alpha \succ \beta \) with respect to C’

<table>
<thead>
<tr>
<th></th>
<th>C’</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ( \alpha )</td>
<td>[**]</td>
</tr>
<tr>
<td>b. ( \beta )</td>
<td>[**] *</td>
</tr>
</tbody>
</table>

In this case, \( \alpha \succ \beta \) with respect to the non-binary constraint C’ because after deleting all the common marks, \( \beta \) contains a mark that \( \alpha \) does not contain. In other words, \( \beta \) contains a proper superset of the marks of \( \alpha \).

In the case of constraints in a strict dominance hierarchy, harmonic orderings are established on the basis of minimal violation of constraints. This means that violation of less dominant constraints is more harmonic than violation of more dominant constraints. In a constraint hierarchy where C \( \gg \) B (A dominates B), and the marks assigned to candidates \( \alpha \) and \( \beta \) are not identical, i.e. not shared, then the candidate with the violation of the lowest ranked constraint is the most harmonic in the ordering. A simple case of constraint dominance is illustrated in tableau (8) below:

(8) \( \alpha \succ \beta \) with respect to C \( \gg \) B

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ( \alpha )</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>b. ( \beta )</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>
Candidate \( \alpha \) is more harmonic, i.e. less marked, than \( \beta \) because it best satisfies the hierarchy \( C >> B \) due to its minimal violation of a lower ranked constraint. If, however, \( \alpha \) also violates constraint \( C \), then \( \alpha \) and \( \beta \) incur identical marks with respect to \( C \). Thus, by the Cancellation Lemma, harmony evaluation is determined solely on the basis of the marks of the lower ranked constraint \( B \), as the following tableau shows:

(9) \( \alpha < \beta \) with respect to \( C >> B \)

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ( \alpha )</td>
<td>[*]</td>
<td>*</td>
</tr>
<tr>
<td>b. ( \beta )</td>
<td>[*]</td>
<td></td>
</tr>
</tbody>
</table>

In (9) \( \alpha \) is less harmonic (we use the symbol "\(<\)"), than \( \beta \), because \( \alpha \) contains a violation of \( B \) that \( \beta \) does not contain. By canceling the common marks assigned by \( C \), candidates \( \alpha \) and \( \beta \) are evaluated by the single binary constraint \( B \), which is only violated by \( \alpha \). Note, however, that if the constraints are ranked in such a way that \( B >> C \), \( \beta \) is still more harmonic than \( \alpha \), since \( \alpha \) violates the dominant constraint \( B \). The violations assigned by \( C \) are no longer relevant in assessing the relative harmonies of \( \alpha \) and \( \beta \) due to the principle of constraint dominance by which constraints higher in the hierarchy have absolute priority over constraints lower in the hierarchy. This situation is illustrated in tableau (10) below:
\((10)\) \(\alpha < \beta\) with respect to \(\text{B} >> \text{C}\)

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. (\alpha)</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>b. (\beta)</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

The fact that in (9) and (10), the ranking of the constraints does not matter in evaluating the relative harmony of these two candidates, is because constraints \(\text{B}\) and \(\text{C}\) stand in a stringency relationship (Prince 1997). Two constraints are in a stringency relationship when they are in a special to general relationship and they disagree on some candidate set, but not conflict, with the general assigning a proper superset of the marks assigned by the specific constraint. In tableaux (9) and (10), candidate \(\beta\) contains a proper subset of the violations of \(\alpha\), which results in \(\beta\) being more harmonic than \(\alpha\) by the basic evaluation procedure of Optimality Theory, i.e. the Cancellation/Domination Lemma, which includes both the Cancellation Lemma and the principle of constraint dominance (Prince and Smolensky, 1993). The lemma is formulated as follows:

\((11)\) Cancellation/Domination Lemma

In order to show that one parse \(\text{B}\) is more harmonic than a competitor \(\text{A}\) which does not incur an identical set of marks, it suffices to show that every mark incurred by \(\text{B}\) is either (i) cancelled by an identical mark incurred by \(\text{A}\), or (ii) dominated by a higher-ranking mark incurred by \(\text{A}\). That is, for every constraint violated by the more harmonic form \(\text{B}\), the losing competitor \(\text{A}\) either (i) matches the violations exactly, or (ii) violates a constraint ranked higher.
Implicit in the *Cancellation Lemma* is a strategy to formalize universal markedness relationships among forms. This strategy does not derive universal markedness relations through fixed universal rankings, as in the case of the place hierarchy in Prince and Smolensky (1993), where the unmarkedness of the place coronal follows from the fixed ranking in (12):

(12)  \[ *_{PL/Lab, *_{PL/Dor} >> *_{PL/Cor}. \]

On the contrary, candidates are evaluated against sets of constraints that are in a stringency relationship. Harmonic orderings of forms are hence established only on the basis of the *Cancellation Lemma*, which involves elimination of the shared marks and evaluation on the basis of the unshared ones. I will call this strategy of analysis to determine universal harmonic orderings the *Subset Strategy*\(^1\) and give it the following formal definition:

(13)  \[ S_1 \lessdot_{UG} S_2 \text{ iff the marks of } S_2 \subset \text{marks of } S_1. \]

A Structure \( S_1 \) is universally less harmonic, and hence more marked, than a Structure \( S_2 \) if and only if the list of marks assigned to \( S_2 \) is a proper subset of the list of marks assigned to \( S_1 \).

In other words, the set of marks assigned to \( S_1 \) contains all of the marks assigned to \( S_2 \) plus one extra mark, which is not assigned to \( S_2 \). A simple example of this strategy can be constructed to evaluate and determine the markedness relationship between a VC structure and a VCC structure. A VCC structure is assumed to be

\(^1\) See the Method of Universal Tableau in Prince and Smolensky (1993) for an illustration of universal rather than language-particular constraint interactions.
more marked than a VC structure. In OT terms, this relationship is expressed as the following harmonic ordering:

\[(14) \quad VC \succ VCC\]

The ordering in (14) can be easily characterized by evaluating the two structures against the relevant constraints on syllable structure, i.e. NOCODA, which bans syllable codas, and *COMPLEX, which bans complex structures. Tableau (15) below demonstrates the harmonic ordering between the two structures.

\[
\begin{array}{|c|c|}
\hline
\text{NOCODA} & \text{*COMPLEX} \\
\hline
a. VC & * \\
\hline
b. VCC & * \\
\hline
\end{array}
\]

As (15) shows, candidate (a), i.e. the VC structure, is universally less marked, and hence more harmonic than a structure VCC, because it contains a proper subset of the marks that VCC contains. Thus the universal markedness relationship between these two syllable structures follows directly from simple comparison of shared and unshared marks, rather than priority of marks derived via fixed rankings.

1.3.3 Implicational Universals

A theory of markedness constructed in this way, together with the basic architecture of Optimality Theory, that centers around constraint interaction as an
explanatory method of analysis, makes it possible to provide a formal characterization of Implicational Universals.

Implicational universals are involved in many typological generalizations and specify that the presence of one structure in a language’s inventory implies the presence of another structure but not vice-versa. Implicational universals of this type are often explained in terms of markedness, in the sense that a marked structure is found in a language only if its unmarked counterpart also occurs. However, markedness relations are established on the basis of the implicational universal observed, thus giving rise to a problem of circularity.

In Optimality Theory implicational universals follow directly from the architecture of the theory. Prince and Smolensky (1993) characterize implicational universals as follows:

(16) An implicational universal of the form ‘ψ in an inventory implies ϕ in the inventory’ holds if, for every possible grammar in which there is some input whose optimal parse includes ψ, there is an input whose optimal parse in that same grammar includes ϕ.

They further formulate the following general strategy to establish implicational universals:

(17) General Strategy for Establishing Implicational Universals ψ ⇒ ϕ

If a configuration ψ is in the inventory of a grammar G, then there must be some input Iψ such that ψ appears in the corresponding output, which, being the optimal parse, must be more harmonic than all competitors. Consideration of some
competitors shows that this can only happen if the constraint hierarchy defining the grammar G meets certain domination conditions. These conditions entail - typically by dint of universal dominations - that an output parse containing $\varphi$ (for some input $I\varphi$) is also optimal.

To clarify how implicational universals are derived in OT, consider the two forms VC and VCC. A syllable of the type VCC violates both NOCODA and *COMPLEX. Admitting VCC into the syllable’s inventory of the language implies that Faithfulness dominates both NOCODA and *COMPLEX, as shown in tableau (18)

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
/VCC/ & Faithfulness & *COMPLEX & NOCODA \\
\hline
\hline
a. $\neq\varphi$ & * & * & * \\
\hline
b. VC & * & ! & * \\
\hline
\end{tabular}
\end{table}

However, the sub-ranking Faithfulness $>>$ NOCODA implies that also VC structures are part of the language’s inventory, as shown in tableau (19) below:

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
/VC/ & Faithfulness & NOCODA \\
\hline
\hline
a. $\neq\varphi$ & VC & * \\
\hline
b. V & ! & \\
\hline
c. V.CV & ! & \\
\hline
\end{tabular}
\end{table}

As a matter of fact, if NOCODA dominated Faithfulness, codas would not be possible in the language at all as the following tableau shows:
Admitting complex codas, therefore, implies admitting simple codas as well. Given these entailment considerations, no ranking of these constraints will ever give a language in which complex codas, but not simple codas, are admitted. The technique illustrated above is called the *Technique of Necessary and Sufficient Conditions*\(^2\) and will be used in the rest of the dissertation to derive the implicational universals holding for clusters.

After having laid out the theory of markedness in OT and the procedures of analysis used in the rest of this dissertation, I turn to a discussion of the Sonority Sequencing Principle within Optimality Theory. The next section is intended as a background to the phenomenon that I will focus on in the rest of the dissertation, i.e. obstruent clusters. Although this dissertation is not about the Sonority Sequencing Principle, a discussion of the basic assumptions that I make about cluster phonotactics in general is necessary in order to understand why obstruent clusters are a unique phenomenon.

\(^2\) This technique is discussed in footnote 72 of Prince and Smolensky (1993) and is also used in Legendre, Raymond & Smolensky (1993).

(20)

<table>
<thead>
<tr>
<th>/VCC/</th>
<th>NOCODA</th>
<th>Faithfulness</th>
<th>*COMPLEX</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. VC</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. VCC</td>
<td>*!</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>c. V</td>
<td></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>}/VCC/</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

2 This technique is discussed in footnote 72 of Prince and Smolensky (1993) and is also used in Legendre, Raymond & Smolensky (1993).
1.4 Background Assumptions

For reasons that will be discussed in Chapter 2, I assume a universal sonority scale such as the one in Clements (1990), which only refers to the major classes of segments (O < N < L < G). Under this scale, two-member clusters are classified as in diagram (21). For each column, the sequence on the left of the comma indicates an onset cluster, whereas the one on the right indicates a coda cluster.

(21)

a. Core Clusters  b. Sonority Reversals  c. Sonority Plateaus

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>OG, GO</td>
<td>GO, OG</td>
<td>GG</td>
</tr>
<tr>
<td>OL, LO</td>
<td>LO, OL</td>
<td>LL</td>
</tr>
<tr>
<td>ON, NO</td>
<td>NO, ON</td>
<td>NN</td>
</tr>
<tr>
<td>LG, GL</td>
<td>GL, LG</td>
<td>OO</td>
</tr>
<tr>
<td>NG, GN</td>
<td>GN, NG</td>
<td></td>
</tr>
<tr>
<td>NL, LN</td>
<td>LN, LN</td>
<td></td>
</tr>
</tbody>
</table>

All the clusters in (21a) are classified as core clusters because they show a decrease in sonority towards the syllable margins and thus follow the SSP. The clusters in (21b) are, instead, classified as reversals because the most sonorous segment occurs closer to the syllable margin than to the syllable peak. The clusters in (21c) instead constitute plateaus since there is no difference in sonority between the members of the clusters.

Clements (1990) proposes to evaluate the relative complexity of the clusters listed in (21a), i.e. the core clusters or unmarked clusters, in terms of the Dispersion Principle. The Dispersion Principle is an evaluation metric that
determines the relative complexity of syllable types on the basis of their degree of
distance from the optimal syllable, i.e. a syllable with the maximal and most
evenly-distributed rise in sonority at the beginning and the minimal drop in
sonority at the end. He also suggests that the relative complexity of reversals and
plateaus may be calculated proportionally to their distance from the unmarked
syllables. In his view, sonority reversals are more complex than sonority plateaus
and the complexity of sonority reversals increases in proportion to the extent of
the reversal.

Whereas Clements’ formalism represents one of the most insightful
approaches to the more unmarked phonotactics, it leaves some of the marked
phonotactics unexplained. Due to the fact that both core clusters and sonority
reversals involve a difference in sonority among the members of the clusters, it
seems reasonable to assume that Clements’ complexity metric, or an extension of
it, could be an adequate method of evaluation for both core clusters as well as
reversals. However, in the case of plateaus, which do not involve a rise or fall in
sonority between the members of the cluster, the distance in sonority between the
two members is equal to zero. In Clements’ complexity metric, a zero sonority
distance means that these clusters are infinitely bad with respect to sonority, but it
does not shed any light on the relative complexity within the set of obstruent
clusters themselves. I propose, therefore, that whereas the relative well-

\footnote{For an OT derivation of Clements’ complexity metric for unmarked syllable types see Smolensky (1995) and Hironymous (1999)}
formedness of core clusters and sonority reversals may be evaluated in terms of sonority by means of the same evaluation procedure, sonority plateaus are different. They need to be explained in sonority-independent terms, because there is no sonority difference between the members of the clusters. The difference must therefore be derived by means of some other parameter. Part of this dissertation is, therefore, devoted to try to fill the gap in the theoretical machinery and provide a formalism to evaluate the relative well-formedness of sonority plateaus.

In this dissertation I focus on obstruent clusters. Obstruent clusters constitute an intriguing phenomenon because of their complex phonotactics. Moreover, the fact that, in a number of languages, obstruent clusters behave differently than core clusters with respect to phonological processes such as syllabification or reduplication, raises the question of what it is that drives such a phenomenon.

1.5 The Sonority Sequencing Principle and Optimality Theory

I formulate the Sonority Sequencing Principle as a positive markedness constraint defining the preferred order of segments within the syllable in the following way:

(22) Sonority Sequencing Principle (SSP)
    Sonority increases towards the syllable peak and decreases towards the syllable margins
Within Optimality Theory, the universality of the SSP and its violability is resolved given the premise that in OT all constraints are in principle violable. OT grammars are constructed in terms of violable constraints and surface patterns are derived via constraint interaction between two basic types of constraints, markedness and faithfulness constraints. Violations of the SSP result from the fact that the SSP, a markedness constraint, is dominated by a faithfulness constraint that requires preservation of input clusters, as I show later. The SSP is therefore not an absolute universal. Absolute universals correspond formally to constraints that are never dominated and therefore never violated.

Moreover, in Optimality Theory, there is only one possible level at which the SSP holds, the level of the output. Unlike lexical phonology which recognized multiple levels of derivation, there are only two levels of representation in OT, the input and the output level, and constraints are stated over output forms only, never on inputs.

Candidates are evaluated for harmony with respect to the SSP following the procedures outlined in the previous section. Evaluation of possible clusters with respect to the single constraint SSP is given in tableau (23) below:

(23)

<table>
<thead>
<tr>
<th></th>
<th>SSP</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. trV</td>
<td></td>
</tr>
<tr>
<td>b. rtV</td>
<td>*</td>
</tr>
<tr>
<td>c. stV</td>
<td>*</td>
</tr>
</tbody>
</table>
Candidate (a) is an example of an onset cluster obeying sonority generalizations. Both (b) and (c) are examples of clusters that violate sonority generalizations. In particular, in candidate (b), the least sonorous segment in the cluster occurs closer to the syllable peak than the most sonorous one. This is an example of a sonority reversal. Candidate (c) is, instead, an example of a sonority plateau, i.e. a cluster in which there is no difference in sonority between the members of the cluster, under the assumption that fricatives and stops form a single class with respect to sonority. Basically, the SSP constraint in an OT grammar has the same role as Clements' version of the SSP, i.e. the Core Syllabification Principle. It classifies clusters into two types, those that conform to the SSP and those that violate it. Formally, candidate (a) is the most harmonic with respect to the SSP constraint because it does not contain the mark that both candidates (b) and (c) contain. In markedness terms, this means that core clusters are the unmarked cluster types, i.e. they satisfy the SSP, and both sonority plateaus and reversals are instead marked with respect to the SSP constraint because they violate it.

The cross-linguistic fact that implicational universals hold between core clusters and clusters that violate the SSP is captured directly from the interaction of Faithfulness with the SSP. If the SSP dominates Faithfulness, only core

---

4 Steriade (1994) argues that clusters of the form s+stop, which are analyzed as plateaus in this dissertation, can actually occur in languages independently of core clusters. In the languages that I surveyed, languages that allow s+stop clusters also allow s+sonorant as in the case of Misantla Totonac (MacKay 1994) or Chiquihuitlán Mazateco (Jamieson 1977; Steriade 1994). I will therefore assume that the occurrence of plateaus and reversals is, in fact, dependent on the occurrence of core clusters.
clusters are allowed to surface because they are the only ones that satisfy the dominant SSP. This is shown in the following tableaux:

(24)

<table>
<thead>
<tr>
<th></th>
<th>SSP</th>
<th>Faithfulness</th>
</tr>
</thead>
<tbody>
<tr>
<td>/trV/</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. ❍ trV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. rtV</td>
<td>*!</td>
<td>*</td>
</tr>
</tbody>
</table>

In tableau (24), an input containing a core cluster surfaces, despite low ranking faithfulness, because it satisfies the dominant SSP. This ranking only allows core clusters to surface in a grammar. An input of the form /rtV/, which is not a core cluster, will never be able to surface faithfully because of the violation of higher ranked SSP, as illustrated in (25) below.

(25)

<table>
<thead>
<tr>
<th></th>
<th>SSP</th>
<th>Faithfulness</th>
</tr>
</thead>
<tbody>
<tr>
<td>/rtV/</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. ❍ trV</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>b. rtV</td>
<td>*!</td>
<td></td>
</tr>
</tbody>
</table>

For an input of the form /rtV/ to surface it is necessary that Faithfulness dominate the SSP, as shown in (26):

(26)

<table>
<thead>
<tr>
<th></th>
<th>Faithfulness</th>
<th>SSP</th>
</tr>
</thead>
<tbody>
<tr>
<td>/rtV/</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. trV</td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>b. ❍ rtV</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>
In the same grammar, an input with a core cluster /trV/ surfaces faithfully as well as well, as in (27)

\[(27)\]

\[
\begin{array}{|c|c|c|}
\hline
/trV/ & Faithfulness & SSP \\
\hline
a. rtV & *! & * \\
\hline
b. trV & & \\
\hline
\end{array}
\]

The implicational universal of the type formulated in Greenberg (1978) which states that clusters violating the SSP always imply the presence of core clusters follows directly from the constraint rankings. If Faithfulness dominates the SSP then both types of clusters are allowed to surface. If the SSP dominates Faithfulness only core clusters are allowed to surface because they are the only harmonic clusters with respect to the markedness constraint SSP.

The SSP constraint, as stated, can only evaluate whether a cluster is well-formed or ill-formed with respect to the sonority generalization expressed by the constraint. The constraint, however, does not say anything about the relative harmony of the various core clusters, nor can it distinguish between the two types of violations that candidates (b) and (c) represent in tableau (23). A system that only consisted of the SSP constraint would not be able to distinguish plateau violations from reversal violations. Moreover, such a system would imply that grammars either disallow or admit any type of sonority violations. This is not a good result, because certain violations are more common than others, and the presence of one type of violation does not necessarily imply the presence of the
other type. For this reason, I believe that the SSP can be best understood as a portmanteau constraint for a whole family of phonotactic constraints. As a first attempt, I will assume that the SSP is actually two separate constraints, which are most likely portmanteau constraints themselves. The two constraints are formulated as negative markedness constraints and, in their simplest form, they ban plateaus and reversals as follows:

(28)  **Plateau**

Sonority plateaus are disallowed

(29)  **Reversal**

Sonority reversals are disallowed

This system of constraints can now formally distinguish the three types of clusters on the basis of their constraint violations, as shown in the following tableau:

<table>
<thead>
<tr>
<th></th>
<th>*Reversal</th>
<th>*Plateau</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. trV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. rtV</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>c. stV</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

Tableau (30) demonstrates the unmarkedness of core clusters, as opposed to the markedness of sonority violations. Sonority plateaus and reversals are less harmonic than core clusters due to their violations of the sonority constraints. Candidate (b), a sonority reversal and candidate (c), a sonority plateau, both contain marks that are not included in the set of the marks assigned to candidate (a), a core cluster. This latter has no marks at all.
The system of constraints proposed for the characterization of cluster phonotactics claims no markedness relationship between plateaus and reversals. This is captured in the fact that the two candidates violate different constraints. The cluster representing a sonority reversal violates *Reversal, whereas the candidate representing sonority plateaus violates *Plateau. Since the two candidates do not share violations with respect to these two constraints, no universal harmonic orderings are established for the two types of clusters. Consequently, no markedness relationships are established between the two types.

Implications exist between core clusters and either plateaus or reversals. These implicational universals follow from the fact that admitting either cluster type in a language will always involve also admitting the more harmonic clusters. Since the candidate containing a core cluster is unmarked with respect to all of the markedness constraints in this system, no matter where faithfulness is ranked, an input containing such a cluster will always surface, regardless of the ranking. For either sonority plateaus or reversals to be admitted in a language, it is necessary that Faithfulness dominates *Plateau or *Reversal, respectively. The rankings Faith >> *Plateau and Faith >> *Reversal do not imply each other and, therefore, no implications between the two types of clusters exist.

1.6 Summary of the chapter.

In this chapter, I have presented an overview of the main issues related to the theory of sonority. I have addressed one of the most basic problems in syllable
phonotactics, i.e. the problem of obstruent clusters and their relation to sonority-based generalizations. I have argued that, given the fact that sonority does not distinguish among these clusters, an insightful understanding of the relevant phonotactics can only be gained by searching for an explanation of their behavior outside of sonority.

In this chapter, I have also presented a discussion of some of the tools that will be used in the analysis, and in particular I have provided a formal characterization of a strategy of analysis that will be used extensively in the rest of the dissertation, i.e. the *Subset Strategy*.

Finally, this chapter has provided an extensive discussion of the Sonority Sequencing Principle and addressed some of the problems that such a principle raises in phonological theories in which constraints are not violable.
2.1 Introduction

The study of obstruent clusters to be presented in the following chapters has two main objectives. On the one hand, it intends to provide new empirical information that can contribute to an understanding of the principles that govern the phonotactics of obstruent clusters in onset position. On the other, by understanding such principles, the dissertation aims to present a new and original analysis of one particular type of obstruent clusters, i.e. s+STOP.

In this chapter, I provide a discussion of the methodology used in the data collection. I then turn to the generalizations observed on the manner dimension for obstruent clusters occurring in onset position. This is the dimension where the feature [continuant] is relevant. On this dimension, the relative harmony of each individual type of obstruent cluster as well as their relative well-formedness and co-occurrence restrictions are captured. In this chapter, I argue that the markedness relationships among the different types of clusters are responsible for the co-occurrence restrictions observed across languages.

The second dimension that will be of interest is the place dimension. On this dimension, restrictions on place features within a cluster are captured. Place generalizations are discussed in Chapter 4.
The third dimension that is relevant to obstruent clusters is the dimension of laryngeal features, i.e. voicing, glottalization and aspiration. I will not be concerned with laryngeal features in this dissertation since they are a well understood aspect of obstruent clusters in the linguistic literature (Lombardi 1991, 1995a, 1995b, 1998; Steriade 1997 and references cited therein).

Finally, in this dissertation I only concentrate on obstruent clusters occurring in onset position. A preliminary investigation of obstruent clusters occurring in coda position has shown that this is indeed not as simple a task as for obstruent clusters occurring in onset position. One of the main problems is that codas are, in general, more restrictive syllable positions than onsets. Consequently, languages that allow obstruent clusters in their coda positions are less common than languages that allow clusters in their onsets. In particular, out of the about 30 languages used in the onset typology, only a very small number (about 5) could have been used for generalizations in the coda. Moreover, codas present the additional problem of weight and extrasyllabicity, which makes a consistent analysis of the data quite difficult. As a matter of fact, two different types of generalizations would be necessary. One type would concern those clusters that are clearly not extrasyllabic in the case of quantity sensitive languages, whereas another would concern quantity insensitive languages in which extrasyllabicity is not an issue. For this reason and due to the scarcity of the data, the present study concentrates solely on onset clusters.
2.2 Methodology

In order to establish consistent criteria for the typological study of obstruent clusters, I have considered the following issues:

- What constitutes an obstruent cluster?
- What is the status of affricates?
- Are certain sequences to be interpreted as single units or clusters?
- How should a representative sample of languages be created?
- Should the generalizations be stated in terms of word or syllable boundaries?
- Are morphologically complex words representative of the language’s phonotactics?
- What about non-native vocabulary?

The following sections will discuss each individual question separately and provide information on the criteria established in each case.

2.2.1 Obstruent Clusters: Definition

Obstruent clusters in this study are defined as tautosyllabic\(^1\) sequences of stops (S) and fricatives (F). The study is restricted to two member sequences, because longer obstruent clusters are much rarer, and it is not always clear whether such sequences constitute examples of minor syllables, i.e. syllables containing a syllabic consonant or a consonant followed by a transitional vowel, or pure

\(^1\) The generalizations are restricted to tautosyllabic clusters. Heterosyllabic clusters may indeed reveal a whole different set of generalizations that are not covered in this study.
consonant clusters.

Obstruent clusters can consist of a fricative and a stop in either order, or of a sequence of two fricatives or two stops. Representative examples of the different types of obstruent clusters are given below:

- **FS** (e.g. English /st/, Havasupai /θp/, Haida /lk/, German /ʃt/)
- **SF** (e.g. Wichita /ks/, Paipai /px/)
- **FF** (e.g. Italian /sf/, Nisqually /sχ/)
- **SS** (e.g. Khasi /pt/, Georgian /tpʰ/)

The obstruent clusters given above represent the four logical ways in which fricatives and stops can cluster. The four possible clusters are all attested across languages.

### 2.2.2 The Status of Affricates

In his study whose focus was to construct generalizations about various types of consonant clusters, Greenberg (1978) considers affricates as clusters of a stop+fricative. Unlike Greenberg, I consider affricates as single segments and, consequently, I don’t consider them instances of SF clusters. As a matter of fact, recent research in feature geometry (Sagey 1986), (Lombardi 1990), has shown that affricates can best be represented as a single root node with two value specifications for the feature [continuant]. Lombardi (1990) presents a number of facts about the affricate that indicate its status as a single segment. In particular, affricates contrast not only with stops and fricatives but also with clusters. For
example, in Polish the cluster [tʃ] contrasts with the affricate [ć] as in the following examples from Campbell (1974):

(1)  trzy  [tʃt]  “three”
czy  [ćt]  “whether”

Affricates also pattern with single segments in syllabification. Chipewyan, for example, only allows simple onsets. Affricates occur in onsets, which suggests that they must be single segments themselves.

Affricates are, moreover, treated as single segments by reduplication processes. For example, in Ewe (Ansre 1963) there is a process of reduplication that copies only the first consonant of a consonant cluster in the root, as shown in example (2a). If the root contains an affricate, the affricate is copied in its entirety, as in examples (2b) and (2c):

(2)  a.  fle  fefle  “buy”
b.  ci  cici  “grow”  *tici
c.  dzra  dzadzra  “sell”  *dadzra

In addition, affricates are never affected by epenthesis or metathesis processes. For example, in Hebrew (Bolozky 1980) the cluster [ts] and the affricate [c] contrast. Whereas the cluster can be broken up by an epenthetic vowel in careful speech, the affricate can never be broken up in the same way:

(3)  /tsumet lev/  [tʃsumet lev]  “attention”
    /cilum/  *[tʃsilum]  “photograph”
Under the view that affricates are single segments, therefore, an affricate does not constitute an obstruent cluster by itself since clusters are, by definition, sequences of two distinct root nodes. Affricates, however, can be one of the members of an obstruent cluster, and thus combine with either fricatives or stops to form a cluster, depending on the language.

2.2.3 SF Sequences: Clusters or Singletons?

For most of the languages considered, the status of affricates was uncontroversial in the sources consulted. However, for the few controversial cases, I decided on the status of certain SF clusters in part on the basis of the facts discovered in the present dissertation for obstruent clusters. For example, in the case of German, where researchers disagree on the status of homorganic \([ts], [tʃ],\) and \([pf]\), I have favored the affricate analysis for these three segments on the basis of the place restrictions observed for true obstruent clusters in the language. In the case of languages with both homorganic as well as non-homorganic SF sequences, such as \([ks], [kʃ], [tf], [kl], [ps]\), I interpret both types of sequences as clusters.

According to Lombardi (1990), whereas tautosyllabic clusters tend not to share place, the two parts of an affricate must share place.

---

2 Assuming Lombardi’s proposal (1990) that the [-cont] and [+cont] components of an affricate are unordered throughout the phonological representation, the system I propose predicts that an affricate can form a clusters with a stop or a fricative depending on what types of obstruent clusters the language allows in the case of simplex segments. For example, in Nisqually in Chapter 3, I show that the ranking that disallows a stop+stop cluster also disallows an affricate+stop cluster.

3 See the analysis of German in Chapter 4.
Along the same line, the choice of treating sequences of SS, SF or FF as clusters rather than single segments in the languages sampled in the study depended mostly on the presence of phonological evidence that would support their status as clusters rather than single segments.

2.2.4 The Sampled Languages

The data was collected from a sample of about 30 languages representative of a number of different language families. Language families represented in the study ranged from Indo-European, to Caucasian, Dravidian, Austroasiatic, Afroasiatic, Tibeto-Burman and Amerindian languages. The Niger-Congo family is represented as well with Eggon. Eggon is indeed an exceptional language in this family because Niger-Congo languages typically disallow consonant clusters. These languages are usually characterized by open syllables. If clusters are allowed, they are generally limited to a few types, such as nasal+obstruent or obstruent+liquid. Eggon, on the other hand, allows obstruent clusters of the FS, SF and SS type (Maddieson 1981).

To make this small sample of languages as representative of the world’s languages as possible, I tried to span across as many language families as possible rather than stay within a few well documented language families. Although concentrating on Indo-European or Amerindian families, for example, would have increased the number of languages in the typology, since obstruent clusters are relatively common across these language families, I don’t believe it would have
contributed to the understanding of universal phonotactics. The languages considered, with their language families are indicated in the following:

<table>
<thead>
<tr>
<th>Language Family</th>
<th>Languages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indo-European</td>
<td></td>
</tr>
<tr>
<td>Germanic:</td>
<td>German, Dutch, English</td>
</tr>
<tr>
<td>Romance:</td>
<td>Italian</td>
</tr>
<tr>
<td>Hellenic:</td>
<td>Greek (Attic, Modern)</td>
</tr>
<tr>
<td>Slavic:</td>
<td>Serbo-Croatian</td>
</tr>
<tr>
<td>Baltic:</td>
<td>Lithuanian</td>
</tr>
<tr>
<td>Indo-Iranian:</td>
<td>Pashto, Hindi</td>
</tr>
<tr>
<td>Caucasian</td>
<td></td>
</tr>
<tr>
<td>Georgian</td>
<td></td>
</tr>
<tr>
<td>Austro-Asiatic</td>
<td></td>
</tr>
<tr>
<td>Mon-Khmer:</td>
<td>Cambodian, Khasi</td>
</tr>
<tr>
<td>Tibeto-Burman</td>
<td></td>
</tr>
<tr>
<td>Qiang:</td>
<td>Mawo</td>
</tr>
<tr>
<td>Tibetan:</td>
<td>Ladakhi</td>
</tr>
<tr>
<td>Dravidian</td>
<td></td>
</tr>
<tr>
<td>Central Dravidian:</td>
<td>Telugu</td>
</tr>
<tr>
<td>Austronesian</td>
<td></td>
</tr>
<tr>
<td>Tsou</td>
<td></td>
</tr>
<tr>
<td>Niger-Congo</td>
<td></td>
</tr>
<tr>
<td>Eggon</td>
<td></td>
</tr>
<tr>
<td>Afro-Asiatic</td>
<td></td>
</tr>
<tr>
<td>Semitic:</td>
<td>Modern Hebrew</td>
</tr>
</tbody>
</table>
2.2.5 Syllable Onsets

Since this study is only concerned with tautosyllabic obstruent sequences, the majority of the data was collected from clusters occurring in word-initial position, but not restricted to them. In some languages, such as Modern Greek, where syllabification of medial clusters is ambiguous, the generalizations observed at the margins proved to hold also in word medial position. For this reason, I have chosen to talk about onset obstruent clusters rather than word initial clusters. Moreover, the arguments provided for the view that obstruent clusters in initial position are extrasyllabic are not always compelling. As I will argue later for Italian, the fact that medial obstruent clusters are syllabified heterosyllabically, does not necessarily mean that word-initially they must be heterosyllabic as well. In other words, I show that the arguments for the heterosyllabicity of word-initial obstruent clusters do not indeed provide evidence for the extrasyllabicity of word initials.
2.2.6 Morphologically Complex Clusters

Most of the data in the study represents generalizations drawn from monomorphemic words. Clusters resulting from morpheme concatenation can provide information about the phonotactics of simplex words. If a language consistently tolerates certain clusters that result from affixation, it may indeed mean that in that language those clusters are well-formed. If they were ill-formed we would expect some phonological process to apply to repair the offending sequence. This is the case of many Tibeto-Burman languages in which a very large number of clusters are derived via affixation. On the other hand, however, one must be wary of the possibility that such clusters may, indeed, be ill-formed in monomorphemic words, but can survive in polymorphemic words for some other independent morphological reason. I argue, in Chapter 5, that this is the case for s+fricative clusters in Italian. Decisions on this issue were made on a language particular basis.

2.2.7 Non-native Phonotactics

Another important issue in the interpretation of the data was to decide whether borrowed words with unusual clusters, or clusters that appeared only in one stratum of the vocabulary of clear foreign origin, should be considered as part of the cluster inventory of the language. Whenever possible, only clusters from words belonging to the native vocabulary were considered as part of the cluster inventory of the language. For example, in languages with clusters present only
in words of clear foreign origin, or in words belonging to a particular lexical stratum of non-native origin, such clusters were considered marginal and thus not necessarily relevant for classificatory purposes. The true properties and universals of cluster inventories can only be captured if a clear distinction is made between native, and therefore productive, and non-native, and therefore non-productive, consonant clusters. Considering clusters that are not part of the native inventory as part of the whole inventory of the language would only contribute to a description of the language rather than lead to an understanding of the universal properties of language.

In the section that follows I present the generalizations that emerge from the cross-linguistic study. The section focuses on the manner dimension in onset clusters. It contributes to the understanding of the principles underlying co-occurrence restrictions of the various types of obstruent clusters.
2.3 Onset Generalizations

The four possible types of obstruent clusters, i.e. FS, FF, SF and SS, give rise to 15 possible ways in which such clusters can either occur in isolation or co-occur in the world’s languages. Table (5) below lists all the logically possible patterns in which obstruent clusters can occur across languages. Of these patterns, only six are shown to occur in the onset in the languages that I have investigated. A check under the onset column indicates that the pattern on the right was found.

(5)

<table>
<thead>
<tr>
<th>Patterns of Occurrence</th>
<th>Onset</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. FS</td>
<td>✓</td>
</tr>
<tr>
<td>2. FF</td>
<td></td>
</tr>
<tr>
<td>3. SF</td>
<td></td>
</tr>
<tr>
<td>4. SS</td>
<td></td>
</tr>
<tr>
<td>5. FS FF</td>
<td>✓</td>
</tr>
<tr>
<td>6. FS SF</td>
<td>✓</td>
</tr>
<tr>
<td>7. FS SS</td>
<td></td>
</tr>
<tr>
<td>8. FF SF</td>
<td></td>
</tr>
<tr>
<td>9. FF SS</td>
<td></td>
</tr>
<tr>
<td>10. SF SS</td>
<td></td>
</tr>
<tr>
<td>11. FS FF SF</td>
<td>✓</td>
</tr>
<tr>
<td>12. FS FF SS</td>
<td></td>
</tr>
<tr>
<td>13. FS SF SS</td>
<td>✓</td>
</tr>
<tr>
<td>14. FF SF SS</td>
<td></td>
</tr>
<tr>
<td>15. FS FF SF SS</td>
<td>✓</td>
</tr>
</tbody>
</table>
As the table shows, there is only a limited number of ways in which these clusters can either occur in isolation or co-occur in the world’s languages. Out of the 15 possible ways in which inventories of onset obstruent clusters of length two can be constructed, only six ways are attested to occur across languages. The following table shows the six different language types\(^4\) and the clusters allowed for each type.

\[
\begin{array}{|c|c|c|c|}
\hline
  & FS & SF & SS & FF \\
\hline
Type 1 & ✓ & & & \\
\hline
Type 2 & ✓ & & ✓ & \\
\hline
Type 3 & ✓ & ✓ & & \\
\hline
Type 4 & ✓ & ✓ & & ✓ \\
\hline
Type 5 & ✓ & ✓ & ✓ & \\
\hline
Type 6 & ✓ & ✓ & ✓ & ✓ \\
\hline
\end{array}
\]

Languages of Type 1 only allow fricatives in initial position and only stops as the second member of the cluster. Examples of Type 1 languages are English (Kenstowicz 1994), Haida (Swanton 1910; Sapir 1922), Havasupai (Seiden 1963; Hinton 1984), Hindi (Nagamma Reddy 1987), Isthmus Zapotec (Marlett and Pickett 1987), Italian (Nespor 1993), Mazateco (Pike and Pike 1947; Steriade 1994), Mislanta Totonac (MacKay 1994), Modern Greek (Joseph and Philippaki-Warburton 1987), Telugu (Nagamma Reddy 1987) and Yuchi (Wolff 1948; Crawford 1973). Type 2 languages allow both stops and fricatives to follow an

\(^4\) The exceptions to these generalizations will be discussed in Section 2.5.2
initial fricative. Dutch (De Schutter 1994) belongs to this class of languages. The languages of Type 3 allow both fricative and stop combinations in either order. No combinations of two fricatives or two stops are allowed. An example of a Type 3 language is Wichita (Rood 1975). Type 4 languages allow combinations of fricatives and stops in either order, and sequences of two fricatives as well. There are no clusters containing two adjacent stops. Nisqually (Hoard 1978), Paipai (Joel 1966; Wares 1968) and Pashto (Penzl 1955) belong to this group of languages. Unlike languages of Type 4, Type 5 languages only disallow a sequence of two adjacent fricatives. Fricatives and stops can combine freely without any restriction on the order of occurrence. Type 5 languages are Attic Greek (Steriade 1982), Dakota (Boas and Deloria 1976) and Khasi (Henderson 1976). Finally languages of Type 6 allow all four logical possibilities. There are no restrictions on the relative order of combinations of fricatives and stops as well as on sequences of segments belonging to the same natural class. Georgian (Vogt 1971; Deprez 1988; Chitoran 1994), Seri (Marlett 1981, 1988), Serbo-Croatian (Hodge 1946), Tsou (Wright 1996) and Yateé Zapotec (Jaeger and Van Valin 1982) belong to this group of languages.

The typology above shows that languages which only allow one type of combination always allow a sequence containing a fricative and a stop, in this exact order. FS is the only cluster that can occur in isolation, it is always present and the presence of other types of combinations always implies its presence. The presence of a sequence of two fricatives always implies the presence of FS
sequences, but it seems to be independent of the other two types of clusters, i.e. SF and SS. However, the presence of SF clusters does imply the presence of FS, but does not imply the presence of either FF or SS. SS sequences imply the presence of SF sequences, and consequently the presence of FS clusters. There seems to be no implicational relation between FF and SS clusters, as well as between FF and SF clusters. These implications are schematized in the following diagram:

(7) \[
\begin{array}{c}
SS \\
\downarrow \\
SF \\
\downarrow \\
FF \Rightarrow FS
\end{array}
\]

In figure (7), implications are shown to exist between SS and SF, SF and FS and by transitivity SS and FS. Assuming implications as a means to determine markedness, the following markedness relations can be established

(8) \[
\begin{array}{c}
FS \Rightarrow FF \\
\downarrow \\
SF \\
\downarrow \\
SS
\end{array}
\]

Diagram (8) shows the markedness relations among the four types of clusters, with FS being the least marked and SS being the most marked given the fact that its presence not only implies the presence of FS clusters but also the presence of
SF clusters. The diagram also shows no relation between FF clusters and SF or SS clusters. The existence of markedness relations derived by implicational universals suggests that any analysis of this kind of clusters must be able to provide a principled account of such an issue.

In the next section I will briefly discuss the inadequacy of sonority as a parameter to account for the generalizations represented here. I argue that sonority is not relevant to obstruent clusters since it fails to account for both markedness relations and implicational universals of onset obstruent clusters.

2.4 Sonority and the Typology

Let us first consider a scale in which obstruents are broken down into stops and fricatives, with stops being less sonorous than fricatives as commonly assumed:

\[(9) \quad F > S\]

Given this scale, the SSP would predict the well-formedness of SF clusters and the ill-formedness of FS clusters with respect to sonority. In other words, it would predict that SF clusters should not only be quite common, or at least more common than FS clusters, which would constitute a violation of the SSP, but also that SF should be the unmarked case along the sonority dimension. Thus, we would expect to find both languages with only SF clusters, as well as languages where the following implication holds:

\[(10) \quad FS \Rightarrow SF\]

If a language has FS clusters, then it has SF clusters.
However, as the typology shows, there are no languages which behave in this way. On the contrary, FS clusters, but not SF clusters, can be found in isolation, and SF always implies the presence of FS, thus making (10) false. As discussed in section 2.2, I argue that FS and not SF is unmarked, which is a reasonable conclusion given the fact that FS clusters are quite common across languages, much more common than SF clusters. A scale which assigns a higher sonority rank to fricatives is, therefore, highly problematic for an account of the typology of onset obstruent clusters and its implicational universals under a sonority-based approach.

Let us explore now the possibility of a scale opposite to (11), i.e. a scale in which stops would be more sonorous than fricatives:

(11) $S > F$

Although there is no independent evidence for such a scale, the typology presented in this dissertation would motivate it, since it would allow us to predict some of the generalizations observed. In other words, under this scale the existence of languages with only FS clusters but not SF and the implication $SF \Rightarrow FS$ would be completely predictable. However, such a scale would still be unable to explain other facts about obstruent clusters. In particular, this scale would be unable to explain the implications $FF \Rightarrow FS$ and $SS \Rightarrow SF$. Sonority cannot therefore be invoked to account exhaustively for the generalizations which emerge from the typology of obstruent clusters.
2.5 Analysis of the Generalizations

In the sections that follow, I will analyze these facts at two separate levels of abstraction. I will first discuss the markedness relations that hold among the four types of clusters and how we can formally derive these relations in OT. I will then discuss the factorial typology and the implicational universals that follow from the constraints proposed to account for the markedness relations among the different types of obstruent clusters. Finally, I will provide a discussion of the Harmonic Bounding Argument and show how this procedure of analysis can account for what clusters count as well-formed or ill-formed under a particular constraint ranking.

2.5.1 Harmonic Orderings or Markedness Relations

The markedness relations schematized in (8) in section 2.3 directly translate to the following harmonic orderings:

(12)  
   a. FS \succ FF  
   b. FS \succ SF \succ SS

To establish the orderings in (10), I propose the following set of markedness constraints:

(13)  OCP[-cont]  
      Tautosyllabic [-continuant] segments are disallowed.

(14)  OCP[+cont]  
      Tautosyllabic [+continuant] segments are disallowed
(15) *SO

A tautosyllabic sequence containing a stop followed by any obstruent is disallowed.

Constraint (15) is a negative constraint, which disallows tautosyllabic sequences of a stop and any obstruent, either a fricative or a stop. It is justified both phonetically and phonologically. Phonetically, it reflects the preference for stops to be released into more sonorous segments. Phonologically, it allows us to assign SS clusters a proper superset of the marks assigned to SF clusters and thus derive the ordering SF ≻ SS. A similar constraint is proposed in Steriade (1994). This constraint will prove crucial in the analysis of obstruent clusters in Modern Greek.

Constraints (13) and (14) are two separate OCP constraints (Leben 1973; Goldsmith 1979; McCarthy 1986; Yip 1988; Odden 1988). They are formulated over each value of the feature [continuant] and state, respectively, that SS or FF sequences are disallowed. The two OCP constraints, as well as the sequential markedness constraint in (15), apply within tautosyllabic clusters. The reason for specifying the domain of application of these constraints lies in the fact that the generalizations that hold for tautosyllabic obstruent clusters may not necessarily hold for heterosyllabic clusters as well.

The relative harmony of the four different types of obstruent clusters is obtained by evaluating them against the three structural constraints given above. Evaluation of the different obstruent clusters is given in tableau (16). The three structural constraints are universally unranked with respect to each other and the relative harmony of the different clusters is obtained via the strategy of analysis.
to determine universal harmonic orderings introduced in Chapter 1, which I referred to as the *Subset Strategy*.

(16)

<table>
<thead>
<tr>
<th></th>
<th>OCP[+cont]</th>
<th>*SO</th>
<th>OCP[-cont]</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. FS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. FF</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. SF</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>d. SS</td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

Along the dimension of onset obstruent combinations, FS is the most harmonic of all the cluster types with respect to this constraint system because it receives no marks at all. FS is provably the unmarked cluster type along the dimension of obstruent clusters. FF and SF are less harmonic, and hence more marked, than FS clusters because both clusters are assigned a mark that FS does not receive. In particular, FF is marked with respect to OCP[+cont] and SF is marked with respect to *SO. The marks that FF and SF receive are not identical, therefore there is no harmonic ordering between the two clusters. In this respect, harmony differs from markedness. Whereas the two clusters, FF and SF, can be said to be equally marked because they both imply the least marked cluster FS, they however cannot be said to be equally harmonic because they do not receive identical marks. No relative harmony can therefore be established between FF and SF. Harmonic ordering, on the contrary, exists between SF and SS. SS is less harmonic than SF because the list of marks of SS includes all of the marks assigned to SF plus one,
i.e. the mark assigned by OCP[-cont]. Note that, since the list of marks of FS is empty, than FS is obviously more harmonic than SS, as well as SF.

### 2.5.2 The Factorial Typology and the Implicational Universals

By interleaving the markedness constraints proposed in the previous section (OCP[+cont], OCP[-cont], *SO) with Faithfulness, the full typology of onset obstruent clusters is obtained as well as its implicational universal. The following table provides a unified picture that illustrates the re-ranking of the constraints in the six different grammars.

(17)

<table>
<thead>
<tr>
<th>LANGUAGE TYPES</th>
<th>CONSTRAINT RANKINGS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type 1</strong>: FS</td>
<td>OCP[+cont] OCP[-cont] *SO &gt;&gt; Faith</td>
</tr>
<tr>
<td><strong>Type 2</strong>: FS-FF</td>
<td>_____ OCP[-cont] *SO &gt;&gt; Faith &gt;&gt; OCP[+cont]</td>
</tr>
<tr>
<td><strong>Type 3</strong>: FS-SF</td>
<td>OCP[+cont] OCP[-cont] ___ &gt;&gt; Faith &gt;&gt; *SO</td>
</tr>
<tr>
<td><strong>Type 4</strong>: FS-SF-FF</td>
<td>_____ OCP[-cont] ___ &gt;&gt; Faith &gt;&gt; OCP[+cont] *SO</td>
</tr>
<tr>
<td><strong>Type 5</strong>: FS-SF-SS</td>
<td>OCP[+cont] ___ ___ &gt;&gt; Faith &gt;&gt; OCP[-cont] *SO</td>
</tr>
<tr>
<td><strong>Type 6</strong>: FS-SF-FF-SS</td>
<td>Faith &gt;&gt; OCP[+cont] OCP[-cont] *SO</td>
</tr>
</tbody>
</table>

The ranking for **Type 1**, where Faithfulness is dominated by the three structural constraints, allows only FS clusters to surface. FS is the unmarked cluster with respect to all structural constraints, therefore whatever ranking is established, it will always surface. However, in order to prevent inputs containing
ill-formed clusters to surface it is necessary that the structural constraints dominate Faithfulness.

Type 2 languages allow FS as well as FF clusters. FS will surface regardless of the ranking, given its unmarked status. However, in order to allow FF clusters in a language it is necessary that OCP[+cont] be ranked below Faithfulness. OCP[-cont] and *SO must dominate Faithfulness to assure that inputs of the form SF and SS do not surface.

Type 3 languages allow FS and SF sequences. Once again, FS will surface regardless of the ranking. For SF to surface it is necessary that *SO be ranked below Faithfulness. OCP[-cont] and OCP[+cont] must be ranked above Faithfulness to avoid that inputs of the form SS and FF can surface.

For FS, SF and FF to surface in languages of Type 4, *SO as well as OCP[+cont] must be ranked below Faithfulness. OCP[-cont] must dominate Faithfulness to prevent an input of the form SS to surface in the language.

In Type 5 languages, *SO and OCP[-cont] must both be ranked below Faithfulness in order to admit SF and SS clusters together with the unmarked cluster FS. In this languages, FF clusters do not surface given that OCP[+cont] dominates Faithfulness.

Finally, for all four cluster types to surface in a grammar it is necessary that Faithfulness be ranked above the three structural constraints. This ranking assures that all four cluster types can surface faithfully in the grammar and thus form well-formed clusters.
Using the *Technique of Necessary and Sufficient Conditions* (Prince and Smolensky 1993), the implicational universals observed in the typology follow directly from entailment considerations on the rankings established to admit the relevant clusters in the inventories of the typological languages. First consider the cluster FS, this is unmarked with respect to all constraints in the hierarchy, therefore whatever ranking is established it will always have an optimal output parse. As for the cluster FF, the necessary and sufficient condition that allows it to surface in a grammar is that Faith >> OCP[+cont]. This ranking, however, also entails that FS will surface given its unmarked status. To allow SF in a grammar, instead, it is necessary that Faith >> *SO. This ranking entails that FS will also surface, but does not entail that FF will surface, as expected given the fact that there is no implication holding between SF and FF. Finally, for SS to be admitted in a grammar it is necessary that Faith >> OCP[-cont], *SO. However this ranking entails that SF will also be admitted in the same grammar, since the ranking Faith >> OCP[-cont], *SO entails the ranking Faith >> *SO. The ranking established for SS therefore assures that the same grammar admits SS as well as SF. In other words, given these logical entailments, there is no grammar that allows SS but not SF or FS, or FF but not FS. The system proposed in this dissertation can never give rise to a language in which the implications in (7) do not hold. In other words, the constraint system proposed will admit only harmonically complete languages.
2.5.3 Harmonic Completeness

According to Prince and Smolensky (1993), “harmonic completeness means that when a language admits forms that are marked along some dimension, it will also admit all the forms that are less marked along that dimension”. The constraint system proposed, thus, only admits harmonically complete languages. This is to say that the typology defined by the constraints proposed has the Strong Harmonic Completeness property. However, this is not to say that harmonically incomplete languages are impossible, i.e. languages in which marked structures are admitted without less marked structures being admitted as well. Other factors may, indeed, come into play that give rise to harmonically incomplete languages. In this type of language a more marked structure surfaces because of the constraint system, but a less marked structure cannot surface due to some other constraint that interacts with the system proposed. In particular an harmonically incomplete language may result from the interaction of various dimensions of markedness. For example, we could easily conceive of a language in which there are only stops but not fricatives. If this language allows obstruent clusters, such clusters would only be of the type SS. The language would therefore lack the less marked FS and SF clusters, but allow the more marked SS and be harmonically incomplete. In Chapter 4, I will discuss the case of Takelma and show how the two dimensions of markedness relevant to obstruent clusters can conspire and yield a harmonically incomplete system.
2.5.4 Relative Well-formedness and Harmonic Bounding

In section 2.4.1, I have argued for the existence of harmonic orderings among the different types of clusters and established an evaluation metric for the computation of the relative harmony of each cluster with respect to each other. In section 2.5.2 I have shown how the ranking of these constraints will give rise to the patterns observed and how entailment considerations will account for the implicational relations observed. In this section I will focus on the relative well-formedness and ill-formedness of each cluster in the grammars predicted by the typology. The discussion will be in abstract terms and basically provide an optimality theoretical implementation of certain Morpheme Structure Constraints (MSC) (Kiparsky 1968), i.e. phonotactic constraints on sound sequences in each individual language. In Optimality Theory, MSC can be derived by the Harmonic Bounding Argument.

In OT, showing that a given structure is well-formed in a grammar is a matter of showing that an output parse containing such structure is optimal in the same grammar. The procedure to determine which output parse is optimal is based on the Cancellation/Domination Lemma (Prince and Smolensky 1993) discussed in Chapter 1 and repeated here for convenience.

(18) *Cancellation/Domination Lemma.* In order to show that one parse B is more harmonic than a competitor A which does not incur an identical set of marks, it suffices to show that every mark incurred by B is either canceled by an identical mark incurred by A, or dominated by a higher ranking mark incurred by A. That is, for
every constraint violated by the more harmonic form B, the losing competitor A either matches the violation exactly, or violates a constraint ranked higher.

To show that a structure is ill-formed in a grammar, instead, involves showing that such structure can never be an optimal output because of some other structure that is provably more harmonic and thus prevents it from surfacing. The general technique of analysis developed to account for ill-formedness is called Harmonic Bounding. The technique is defined in Prince and Smolensky (1993) as follows:

(19)  *Harmonic Bounding*. In order to show that a particular structure $\varphi$ does not appear in the outputs of a grammar, it suffices to show that any candidate structure A containing $\varphi$ is less harmonic than *one* competing candidate B not containing $\varphi$ (of the same input). (B provides a harmonic (upper) bound for A).

According to this method, in order to show that a structure $\varphi$ is ill-formed in a given grammar “it is sufficient to show that there is always a B-without-$\varphi$ that is better than any A-with-$\varphi$”. For the Harmonic Bounding argument to be successful, B does not necessarily need to be optimal, it is sufficient to show that B is more harmonic than A. Whether B is optimal is a separate issue. Proving that B is more harmonic than A, it is therefore enough to show that no structure containing $\varphi$ will ever be optimal and that it will never occur in any output of the grammar.
Therefore, in terms of the typology proposed in this dissertation, to say that a cluster is ill-formed in a language is not to say that the cluster is not a possible input, but rather that no output of the grammar ever contains that cluster. A well-formed cluster in a grammar is one which is allowed to surface and corresponds to an optimal output candidate. An ill-formed cluster is, instead, one which, although in the input, is not allowed to surface by the constraint system and can never correspond to an optimal output candidate. To show that a cluster A in a given grammar does not surface, it is sufficient to show that there is one candidate B which is provably more harmonic than A. The cluster A is therefore bounded by the better candidate B (the harmonic (upper) bound). Candidate B, however, although more harmonic than A, may not necessarily be the optimal candidate. Throughout the analysis, I will show not only that for any impossible cluster in a given grammar, there is always a harmonic upper bound, but also that all the possible clusters in the same grammar represent harmonic bounds. Although irrelevant for the Harmonic Bounding Argument, among the harmonic bounds the FS candidate almost always turns out to be the most harmonic. Interestingly, FS is the unmarked obstruent cluster type.

Before illustrating the analysis in one of the typological grammars, a brief discussion of the inputs and the candidate set is necessary to understand what makes a cluster in a certain typological grammar either well-formed or ill-formed.
2.5.5 The Inputs and the Candidate Set

The four logical clustering possibilities, i.e. FS, SF, FF and SS are all considered to be possible inputs in any of the six grammars constructed for the six types of languages, because in OT inputs are universal and cannot be restricted. This principle is referred to as *Richness of the Base* (Prince and Smolensky 1993). The ill-formedness of certain clusters with respect to a given grammar is, therefore, obtained by showing that no input leads to an optimal output that contains such clusters, rather than rejecting them as inputs.

As for the candidate set, the only candidates which need to be considered for well-formedness considerations in the typological grammars are four structures which essentially contain the 4 possible clusters, FS, SF, FF and SS. The four candidates are produced by either changing or maintaining the value for the feature [continuant] on one or both segments of the input sequence as exemplified in the following table:

(20)

<table>
<thead>
<tr>
<th>Candidates</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cand1</td>
<td>both segments faithful to input values for [continuant]</td>
</tr>
<tr>
<td>Cand2</td>
<td>only first segment unfaithful to input value for [continuant]</td>
</tr>
<tr>
<td>Cand3</td>
<td>only second segment unfaithful to input value for [continuant]</td>
</tr>
<tr>
<td>Cand4</td>
<td>both segments unfaithful to input values for [continuant]</td>
</tr>
</tbody>
</table>

Obviously, the four candidates considered do not exhaust the range of possible candidates available to each input. Consider for example an input of the type FS.
There are at least two candidates which will have the structure SF for this input. One is obtained by changing the value for [continuant] on both segments, and the other one is the result of metathesis, i.e. correspondent segments maintain the same value for [continuant] but their linear order is reversed. These two candidates violate two different faithfulness constraints. The former candidate violates a constraint of the Ident(F) family, i.e. \textbf{Ident(cont)}, since both output segments have different input values for the feature [continuant]. The latter, on the contrary, violates \textbf{Linearity} (McCarthy and Prince 1995). This constraint basically says that any two elements of a string stand in an order relation which is necessarily preserved under linearity. Tableau (21) below shows the different violations incurred by the different candidates.

\begin{table}[h]
\centering
\begin{tabular}{|l|c|c|}
\hline
/F_iS_j/ & Ident(cont) & Linearity \\
\hline
a. \text{S}_iF_j & ** & \\
\hline
b. \text{S}_jF_i &  & * \\
\hline
\end{tabular}
\end{table}

As shown in the tableau, candidate (a), which is obtained by changing the input value specification on both segments, incurs two violations of Ident(cont), one for each segment. Candidate (b), instead, does not incur any violation of Ident(cont), since the value of the correspondent segments is not changed as shown by the indexes, but their linear order is. Candidate (b) is a candidate that shows metathesis of the input cluster. In other words, this candidate is an example of a possible repair strategy and is not listed in (20). Since none of the languages that I
surveyed shows this type of repair strategy⁵, such a candidate as well as the relevant constraint will be omitted from tableaux.

Finally, there are also two other candidates which need not be considered for well/ill-formedness evaluations. The two candidates are one candidate in which one of the segments is deleted (deletion candidate), and one candidate whose obstruent sequence is broken up by the insertion of an epenthetic segment (epenthesis candidate). These two candidates exemplify possible “repair strategies”, i.e. strategies a language would adopt to repair a sequence of consonants which is not allowed to surface in a given grammar. These candidates are not relevant candidates to consider for the question of relative well-formedness. This section attempts to account for how typological grammars construct their inventories of obstruent clusters and not how they would repair bad clusters resulting from morphological or phonological processes. In other words, this analysis is concerned with relative well-formedness and ill-formedness and not with phonological alternations. Determining which of the possible repairs strategies a language adopts to avoid ill-formed sequences is an independent question which I will discuss separately later in the chapter. In the next section, I will show how the Harmonic Bounding Argument works in imposing limitations on possible surface clusters in one of the typological grammars, i.e. Type 1

⁵ A brief discussion of possible repair strategies is provided in Section 2.4.6. A typology of repair strategies is, however, outside the scope of this dissertation, because the dissertation focuses on well-formedness rather than repair strategies.
languages.

2.5.6 Harmonic Bounding and Well/ill-formedness in Type 1 languages

Type 1 languages are the languages which only allow FS as onset obstruent clusters. The analysis I will present in this section is about the relative well-/ill-formedness of each type of cluster in languages of this type. The analysis is not about how ill-formed clusters would be repaired but rather why the ill-formed clusters in this type of language can never make it to the surface. I will show that in Type 1 languages no cluster other than FS can ever surface because FS represents the harmonic upper bound. Given any input, the candidate containing this cluster is always more harmonic than any other competing candidate, hence the only cluster allowed to surface. This result is obtained via interleaving the markedness constraints with the faithfulness constraint defined below:

(22) Ident(cont)\(^6\)

Correspondent segments have the same value for the feature continuant.

This constraint belongs to the Ident(F) Constraint Family (McCarthy & Prince, 1995) and assures that input and output segments agree in the specification for the feature [continuant]. The general schema of the constraint is given below:

(23) Ident(F)

Let \(\alpha\) be a segment in S1 and \(\beta\) be any correspondent of \(\alpha\) in S2. If \(\alpha\) is \([\gamma F]\), then \(\beta\) is \([\gamma F]\)

\(^6\) Remember that the point of this analysis is to evaluate relative well-formedness and not repair strategies. This is the main reason for using Ident(cont) rather than any other Faithfulness constraint.
The constraint ranking that determines well-formedness in a *Type 1* language is given in (24):

(24) OCP[+cont], OCP[-cont], *SO >> Ident(cont)

In this grammar, the three structural constraints are unranked with respect to each other but crucially dominate the faithfulness constraint Ident(cont), as shown in the tableaux demonstrating the ill-formedness of the non-occurring clusters. Obeying the requirement on the structure of obstruent clusters is therefore more important than faithfulness to the input in this grammar. In what follows I will first show how the constraint ranking established for *Type 1* languages will make FS surface while disallowing the remaining three cluster types and making them ill-formed in this type of languages.

Let us first start by considering an input of the type FS. The analysis is displayed in tableau (25):

(25)
The FS candidate is the only candidate which does not incur any violation of the structural constraints. It does not violate Ident(cont) either since it is faithful to the input. FF, SF and SS all fail because they all violate one of the high ranked structural constraints. Specifically, the FF candidate incurs a violation of OCP[+cont], and the SF candidate contains a sequence which violates *SO. Finally the SS candidate fails because of both OCP[-cont] and *SO, whose violations are equally fatal since they are unranked with respect to each other. The candidate containing the cluster FS is thus not only more harmonic than all the others, given the fact that it incurs no violations at all, but it is also optimal by virtue of being the most harmonic. This cluster is therefore well-formed with respect to this grammar and hence present in the languages of this type. As can be noticed, the input FS in this grammar does not provide any ranking argument for why Ident(cont) must be at the bottom of the hierarchy. An unranked Ident(cont) would, as a matter of fact, get the same result. It is only in the case of inputs which lead to ill-formed clusters that the ranking in (24) is crucial, as shown below.

With this constraint ranking where the structural constraints dominate faithfulness, given any input, none of the candidates which contain an input that violates the requirements imposed by the structural constraints will ever be optimal. The only candidate that does not violate any of the structural constraints, i.e. FS, will always be more harmonic than any of the competing candidates that instead violate one or more of the structural constraints. FS is therefore the
harmonic bound that ensures that none of the other inputs will ever surface in this grammar. The Harmonic Bounding Argument is illustrated in the following tableaux.

(26)

<table>
<thead>
<tr>
<th>/FF/</th>
<th>OCP[+cont]</th>
<th>OCP[-cont]</th>
<th>*SO</th>
<th>Ident(cont)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. FS</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>b. FF</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. SF</td>
<td></td>
<td>*!</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>d. SS</td>
<td>*(!)</td>
<td>*(!)</td>
<td>**</td>
<td></td>
</tr>
</tbody>
</table>

In (26) the FF candidate is bound by the FS candidate. FS is more harmonic than FF since it incurs a violation of a lower ranked constraint. Ident(cont) must therefore be ranked at the bottom of the hierarchy for FS to be more harmonic than any of the other candidates. If Ident(cont) was unranked, the one mark assigned to FS would be as bad as any of the marks assigned to the competing candidates. FS could not be proven to be better than FF.

(27)

<table>
<thead>
<tr>
<th>/SF/</th>
<th>OCP[+cont]</th>
<th>OCP[-cont]</th>
<th>*SO</th>
<th>Ident(cont)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. FS</td>
<td></td>
<td></td>
<td></td>
<td>**</td>
</tr>
<tr>
<td>b. FF</td>
<td>*!</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>c. SF</td>
<td></td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. SS</td>
<td>*(!)</td>
<td>*(!)</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>
In (27) and (28), the candidate FS provides the harmonic bound which prevents SF and SS, respectively, from surfacing in this grammar. They are therefore, just like FF, ill-formed with respect to the constraint system of Type 1 languages.

### 2.5.7 Ill-formedness and Repair Strategies

The next question to be considered is what happens to ill-formed clusters that may arise in a particular language due to some phonological or morphological process. There are at least three ways in which individual languages can repair ill-formed clusters, although as we will see in the case of Nisqually, these are not the only possible repair strategies. A language can either delete one of the two segments in the cluster, or break up the offending sequence by inserting an epenthetic segment (usually a vowel), or change the ill-formed sequence into a well-formed one, i.e. by neutralization to the unmarked FS, as in the case of Modern Greek.

In (17), I used general Faithfulness to construct the typology of onset obstruent clusters. However, for each language type, there are at least three subtypes of languages based on the strategy that the language adopts to repair ill-
formed clusters. Depending on the repair strategy, Faithfulness in (17) is replaced
by one of the correspondence constraints, MAX-IO, DEP-IO (McCarthy and
Prince 1995) or Ident(cont), which was discussed in 2.4.5. MAX-IO states that
every segment of the input has a correspondent in the output. It prohibits
phonological deletion. DEP-IO states that every segment of the output has a
correspondent in the input. This constraint prohibits phonological epenthesis.

If DEP-IO is ranked in place of Faith in (17), and the other faithfulness
constraints are also ranked higher than DEP-IO, we define a grammar where ill-
formed clusters are repaired by inserting an epenthetic segment. If MAX-IO is
ranked in place of Faith in (17), we define a grammar that repairs ill-formed
clusters by deleting one of the segments. Finally, if Ident(cont) replaces Faith in
(17), and the remaining two faithfulness constraints are higher ranked, we get a
language in which marked clusters are neutralized to the unmarked FS. In what
follows I will discuss each of the three possible repair strategies and show the
relative ranking among the three different faithfulness constraints.

The ranking Ident(cont), MAX-IO >> DEP-IO defines a grammar where
the offending clusters are repaired by epenthesis as shown in tableau (29).
Consider an input SF, which, given the constraint system for Type 1 languages, is
not allowed to surface. DEP-IO being lower ranked with respect to the other
faithfulness constraints will make this input surface as $S\bar{o}F$ (i.e. a sequence
containing an epenthetic segment).
The three candidates shown in tableau (29) include (a) a candidate whose structure contains the only possible cluster in the language (obtained by changing the feature specifications), a deletion (b), and an epenthesis candidate (c). The input cluster will surface as the epenthesis candidate since this candidate incurs a violation of a lower ranked constraint, and hence a lesser violation with respect to the other two candidates.

The ranking Ident(cont), DEP-IO >> MAX-IO defines instead a grammar which repairs ill-formed clusters by deleting one of the segments as shown in tableau (30):

(30)

<table>
<thead>
<tr>
<th>/SF/</th>
<th>Ident(cont)</th>
<th>DEP-IO</th>
<th>MAX-IO</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. FS</td>
<td>**!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. F'</td>
<td></td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>c. S©F</td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

For this discussion, it is irrelevant whether the first or second segment in the cluster is deleted.
Tableau (30) shows that given this constraint ranking, the deletion candidate will turn out to be optimal because it incurs a minimal violation with respect to the other candidates.

Finally the ranking MAX-IO, DEP-IO >> Ident(cont) defines a grammar that neither deletes nor epenthesizes, but rather turns an ill-formed cluster into the unmarked case for obstruent clusters, i.e. it adopts neutralization of marked structures into the unmarked one. For example it turns an SS cluster into an FS, as exemplified in tableau (31):

(31)

<table>
<thead>
<tr>
<th></th>
<th>MAX-IO</th>
<th>DEP-IO</th>
<th>Ident(cont)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. /SS/</td>
<td>FS</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>b. S</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. SoS</td>
<td></td>
<td>*!</td>
<td></td>
</tr>
</tbody>
</table>

To show how exploded Faith interacts with the hierarchy of constraints that I propose, consider for example a language which belongs to Type 3 (i.e. only FS and SF are allowed) and that repairs ill-formed clusters via epenthesis. The constraint hierarchy for such a language would be the one given in (32):

(32) OCP[+cont], OCP[-cont], Ident(cont), MAX-IO >> DEP-IO >> *SO

Note that this predicts that marked clusters can only turn into unmarked FS. In the analysis of Modern Greek in Chapter 3, I will show that a marked cluster, in particular a certain type of FF clusters can actually turn into SF rather than FS under special circumstances.
This hierarchy still assures that a cluster such as SF can surface in this language as shown in tableau (33):

(33)

<table>
<thead>
<tr>
<th>/SF/</th>
<th>OCP[+cont]</th>
<th>OCP[-cont]</th>
<th>Ident(cont)</th>
<th>MAX-IO</th>
<th>DEP-IO</th>
<th>*SO</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. SF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>b. FF</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. FS</td>
<td></td>
<td></td>
<td>**!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. SS</td>
<td></td>
<td>#!</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>f. SaF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>#!</td>
</tr>
</tbody>
</table>

In tableau (33) the SF candidate will surface since it only incurs a violation of the lowest ranked constraint, *SO. At the same time, this ranking assures that a cluster of the form SS in this language does not surface, as the following tableau shows. However, given an input that cannot surface, the addition of MAX-IO and DEP-IO in the hierarchy, shows how this language will repair such an input. Consider for example an input SS as in the following tableau:

(34)

<table>
<thead>
<tr>
<th>/SS/</th>
<th>OCP[+cont]</th>
<th>OCP[-cont]</th>
<th>Ident(cont)</th>
<th>MAX-IO</th>
<th>DEP-IO</th>
<th>*SO</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. SF</td>
<td></td>
<td></td>
<td></td>
<td>#!</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>b. FF</td>
<td>*!</td>
<td></td>
<td>**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. FS</td>
<td></td>
<td></td>
<td>#!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. SS</td>
<td></td>
<td>#!</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>e. S</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f. SaS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>
In tableau (34), the violation incurred by candidate (f) is minimal with respect to the violations of all other candidates. This candidate is optimal. It is therefore the candidate which will surface given an ill-formed input.

Compare tableau (34) with tableau (35) now. In tableau (35) a different relative ranking of the faithfulness sub-hierarchy is exemplified, i.e. Ident(cont), DEP-IO >> MAX-IO. This different relative ranking, interacting with the constraint system proposed, will specify a language which repairs an ill-formed cluster by deleting one of the segments rather than inserting an epenthetic segment.

\[
\begin{array}{|c|c|c|c|c|c|}
\hline
& OCP[+cont] & OCP[-cont] & Ident(cont) & DEP-IO & MAX-IO & *SO \\
\hline
a. SF & & & *! & & & * \\
\hline
b. FF & *! & & & & & \\
\hline
c. FS & & ** & & & & \\
\hline
d. SS & & & *! & & & \\
\hline
e. S & & & & & * & \\
\hline
f. SaS & & & & & *! & \\
\hline
\end{array}
\]

With DEP-IO dominating MAX-IO, inserting an epenthetic segment is a worse violation, in this grammar, than deleting one of the segments.

In the next section I discuss Greenberg’s generalizations about obstruent clusters. I show how the onset generalizations identified in this dissertation improve upon Greenberg’s generalizations.


2.6 Greenberg’s Generalizations

Greenberg (1978) proposes four universals for obstruent clusters. Such universals express the preference for combinations of stops (S) and fricatives (F) as opposed to stop+stop or fricative+fricative in both initial and final systems. The preferences are expressed in terms of implicational universals of the form $\phi \Rightarrow \psi$, whereby the presence of a structure $\phi$ implies the presence of a structure $\psi$, but not vice versa. In initial position, Greenberg formulates the following two universals, respectively 7 and 9 in the original paper:

(36) “In initial systems the presence of at least one combination of stop+stop implies the presence of at least one combination of stop+fricative”.

(37) “In initial systems the existence of at least one fricative+fricative combination implies the presence of at least one stop+fricative combination or at least one fricative+stop combination.”

Using the notation introduced in the previous section, Greenberg’s universals can be represented as follows:

(38) $SS \Rightarrow SF$
(39) $FF \Rightarrow SF, FS$

The implicational universal in (38) is based on Greenberg’s observation that out of 25 languages with SS clusters, all of them also contained SF clusters and two, Huichol and Takelma, did not contain FS clusters. The generalization in (38) thus differs from the generalizations that I propose because, according to Greenberg,
SF⇒FS does not hold. On the other hand, the implication in (39) is based on the fact that out of 33 languages containing FF clusters, only one language contained FF and SF but not FS (Karen) and two contained FF and FS but not SF (Icelandic and Kashmiri). No implications are discussed for FS and SF clusters. The generalization in (39) also differs from the generalizations proposed in (7) because, according to Greenberg, there is an implication between FF and SF that I did not find in my corpus.

The main problem with Greenberg’s generalizations lies in the fact that he counts affricates as stop+fricative clusters. In other words, his generalizations cannot be an adequate representation of the principles of obstruent clusters due to the fact that single segments are confused with clusters. This gives rise to a faulty typology. As a matter of fact, many languages do have affricates in their inventory without necessarily admitting any complex onsets. This explains why in both Greenberg’s universals (5) and (6), SF clusters are always present. Unlike Greenberg, in the universals I propose, the presence of any other obstruent cluster always implies the presence of FS clusters rather than SF. According to Greenberg, however, Huichol and Takelma contain SS and SF but not FS, and Karen contains FF and SF but not FS. These three languages would, as a matter of fact, violate the generalizations I propose because of the absence of FS clusters and the presence of more marked clusters.

None of these three languages, however, constitute a problem for the generalizations proposed in section 2.2. In particular, according to Greenberg,
Huichol contains SF and SS clusters but not FS clusters. But according to the source (McIntosh, 1945), the language actually contains the alveolar and alveopalatal affricates /c, č/, which in Greenberg’s analysis constitute obstruent clusters of the type SF, but in my analysis do not\(^9\). As for the SS clusters, all the clusters of this type occur in morphologically complex words whose initial segment is consistently /p/ or /c/, both of which have morphological content. Given that only SS clusters arising from affixation are found in the language, it can be assumed that obstruent clusters are indeed ill-formed in the language. Their exceptional occurrence can be explained by reference to a constraint that preserves morphological information, as in the case of Italian s+Fricative clusters to be discussed in Chapter 4.

As for Takelma, Greenberg claims there are only SS and SF clusters in the language, but no FS clusters. Again what Greenberg considers an SF cluster is in reality the palatal affricate /ts/, because according to Sapir (1922) the only common initial clusters in Takelma are [tʰp], [tʰk], [sp], [sk], i.e. instances of SS and FS. Under the implication that I have proposed where SS ⇒ SF ⇒ FS, Takelma would represent a potential violation because it contains SS clusters without also allowing SF clusters. However, I will argue in Chapter 4 that Takelma is indeed a *harmonically incomplete* language which allows more marked structures at the expense of less marked ones. In Takelma the two

\(^9\) The arguments for the monosegmental status of affricates are provided in Section 2.2.2.
dimensions that I will discuss in this dissertation, i.e. the manner and place dimensions, come into conflict and give rise to an incomplete system.

Finally, according to Greenberg, Karen contains FF but not FS clusters. Contrary to Greenberg’s claim, none of the Karenic languages for which I have found data allow obstruent clusters. Karen languages are spoken in large areas of Burma and Thailand and only allow core clusters. Based on the data in Kato (1995), Karenic languages do not constitute violations of the generalizations I provide because they do simply not allow for obstruent clusters at all.

To conclude, I have shown, that although it is possible to find languages that violate the typology I propose, it does not necessarily mean that the typology is incorrect. Harmonically incomplete languages are not necessarily a challenge for the implicational relations holding among the four different types of obstruents. Harmonically incomplete systems exist due to independent factors, in each individual language, which interact with the rest of the grammar. In other words, the typology itself has the Property of Strong Harmonic Completeness, i.e. the constraint system gives rise only to harmonically complete languages. Harmonically incomplete languages may, however, result from the interaction of other markedness dimensions.
CHAPTER 3

CASE STUDIES

3.1 Introduction

In this chapter, I present two case studies: Modern Greek and Lushootseed-Nisqually. Both languages represent examples of *harmonically complete* systems. An example of an *harmonically incomplete* language will be provided in Chapter 4.

In the case of Modern Greek, I show a case in which the harmonic upper bound FS for ill-formed clusters actually corresponds to the optimal candidate. In other words, Modern Greek is a language that repairs ill-formed clusters by neutralization to the unmarked structure, i.e. FS clusters. Moreover, I will argue that a unified account of the various phonological processes affecting obstruent clusters is the result of one single constraint ranking established for the language.

The second case study is Lushootseed-Nisqually. This language is interesting for the purposes of the dissertation because it shows a different type of repair strategy for the ill-formed clusters. In particular, in Nisqually, ill-formed obstruent clusters are repaired by obstruent syllabicity; which I argue is completely predictable from simple interactions of syllable structure constraints with the constraints relevant to obstruent clusters.
3.2 Case study I: Modern Greek

3.2.1. Introduction

The analysis of Modern Greek that I present in this section exemplifies a grammar of a *Type 1* language, i.e. a language in which only FS clusters are well-formed. This is despite the fact that FS are not the only clusters that occur in the language. I claim that FS are the only clusters admitted by the constraint ranking defining obstruent clusters. All other types of obstruent clusters that occur in the language are argued to be ill-formed with respect to constraint hierarchy defining obstruent clusters. They, however, are allowed to surface, because they best satisfy independent constraint that have priority over the latter hierarchy.

Moreover, I show that a number of apparently unrelated phonological processes affecting obstruent clusters can be explained as a single process of neutralization to the unmarked FS. In particular, one of the processes affecting obstruent clusters in Modern Greek provides crucial evidence for the activity of the constraint *SO*. The Modern Greek data also provides evidence for recent models of lexicon stratification (Fukazawa 1997, 1999), Fukazawa, Kitahara and Ota (1998) and Itô and Mester (1998).

I will first provide a discussion of the current language situation and a description of the Modern Greek sound system and surface phonotactics patterns.
3.2.2 Lexical Strata in Modern Greek

Modern Greek (Joseph and Philippaki-Warburton (1987), Kaisse (1989)) consists of two separate but co-existing lexical strata: katharevousa and dimotiki. Katharevousa is a sort of artificial archaic language which has mostly been used by conservative administrations and attempts to preserve an older state of the language by borrowing many words from Ancient Greek. Dimotiki represents, instead, the spoken common language. The two strata differ, among other things, in the types of obstruent clusters they allow. Whereas katharevousa permits a richer system of obstruent clusters due to its closer connection with Ancient Greek, in dimotiki voiceless obstruent clusters must disagree in continuancy. The only voiceless obstruent clusters found in the dimotiki vocabulary are, therefore, FS and SF types. When clusters agreeing in continuancy arise in the dimotiki lexicon due to morphological processes or borrowings from classical sources, they generally undergo some type of dissimilation. I will argue that the different processes affecting obstruent clusters in dimotiki are all driven by a single stratum-specific process: neutralization to the unmarked FS.

In the next section, I will briefly review the obstruent system of Modern Greek and provide information on relevant syllabification.

3.2.3 Modern Greek obstruent system and syllable structure

The chart in (1) representing the obstruent phonemes of Modern Greek is based on Joseph and Philippaki-Warburton (1987).
The parentheses around some of the elements in the chart indicate that their phonemic status is not agreed upon by most Greek linguists. As for [ts] and [dz], I follow Householder (1964) and Joseph and Philippaki-Warburton (1987) in analyzing them as simple consonants.

Modern Greek as a whole allows for a variety of clusters both in initial and medial position. Besides core clusters of the form Obstruent+Sonorant, the language also has a rich system of obstruct clusters mainly due to the co-existence of the two sub-lexicons, *katharevousa* and *dimotiki*. The spoken language, depending on the register and social context, may, in fact, contain words of both systems that are no longer perceived as belonging to two separate language forms. For this reason, not only do we find clusters of the type FS and SF, in which the two members differ in continuancy, but also clusters of the type FF and SS, because they are well-formed in the *katharevousa* sub-lexicon (Joseph and Philippaki-Warbuton, 1987). Examples of all the four types of clusters are given below:

<table>
<thead>
<tr>
<th></th>
<th>Bilabial</th>
<th>Inter-dental</th>
<th>Dental</th>
<th>Velar</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stops</strong></td>
<td>Simple</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>p (b)</td>
<td></td>
<td>t (d)</td>
<td>k (g)</td>
</tr>
<tr>
<td><strong>Fricatives</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-sibilant</td>
<td>f v</td>
<td>θ δ</td>
<td></td>
<td>x y</td>
</tr>
<tr>
<td>Sibilant</td>
<td></td>
<td>s z</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Affricates</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sibilant</td>
<td></td>
<td>(ts) (dz)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In general, therefore, given the richness of the system of obstruent clusters, Modern Greek could be classified as a *Type 6* language. However, discussion of the processes affecting obstruent clusters will show that this classification is inaccurate in this case. I will argue, instead, that regardless of the presence of SS, SF and FF clusters, Modern Greek is indeed a *Type 1* language, i.e. a language in which only FS clusters are well-formed with respect to the hierarchy defining obstruent clusters. Within the *dimotiki* stratum, there is crucial evidence that
dimotiki SF\(^1\) clusters are indeed ill-formed with respect to the obstruent cluster constraints. In other words, although a restricted number of SF clusters are allowed to surface in the dimotiki sub-lexicon, I argue that \(^*\)SO dominates faithfulness, thus making SF clusters in general ill-formed with respect to the hierarchy presented in the previous chapter.

Another important feature of Modern Greek is syllabification of medial clusters. According to Setatos (1974) and Joseph and Philippaki-Warburton (1987), Modern Greek tends to follow the Onset Maximization Principle in most circumstances. Obstruent clusters in medial position are mostly syllabified as tautosyllabic onset clusters if the onset is an acceptable word initial cluster. So, for example, the lexical item [efxaristo] (meaning “thank you”) is syllabified as [e.fxa.ri.stó] rather than [ef.xaris.tó]. In case the medial cluster is the result of affixation, however, syllabification may either coincide with the morpheme boundary or follow onset maximization. For example the word [ek+této] ”expose”, may be syllabified as either [ek.té.to] or [e.kté.to]. It is exactly because of languages such as Modern Greek, in which medial obstruent clusters are ambiguously syllabified heterosyllabically, that I have chosen to talk about obstruent clusters in terms of syllable onset and coda rather than word-

\(^1\) FF and SS clusters are also ill-formed with respect to the hierarchy of obstruent cluster constraints in the dimotiki lexicon. Their occurrence in the language will be shown to depend on different principles that the ones that allow certain SF clusters to surface.

\(^2\) According to Setatos (1974) the main reason for preferring the form [e.fxa.ri.stó] is because this word may also occur in casual speech without the initial [e].
initial and final positions. For the sake of simplicity, I assume that all medial obstruent clusters follow the Onset Maximization Principle.

3.2.4 Dimotiki Obstruent Clusters

As discussed previously, Modern Greek consists of two co-existing lexicons which show different patterns of occurrence of obstruent clusters. The katharevousa sub-lexicon allows all four types of obstruent clusters. The obstruent clusters that occur in the dimotiki lexicon are either of the FS or SF type\(^3\). Of these latter, however, only clusters consisting of a stop followed by /s/ are found. In the analysis I propose, I argue that SF clusters are in general ill-formed in the dimotiki sublexicon with respect to the constraints defining the typology of obstruent clusters, just like FF and SS are. The reason for treating SF clusters in general as ill-formed is because they undergo the same neutralization process as FF and SS clusters. However, I will show that, among the SF clusters, STOP+s clusters constitute a privileged subset of SF clusters in this language and are therefore allowed to surface.

In the following sections, I will first argue that SF clusters in general are bad, i.e. ill-formed with respect to the hierarchy defining obstruent clusters, in the dimotiki sublexicon. Their ill-formedness entails that the constraint *SO

---

\(^3\) This restriction may only hold for voiceless obstruent clusters. Voiced FF clusters occur in Modern Greek as well, e.g. [vðelə] leech, [vŋazo] I take out, [zvino] erase. These words most likely belong to the katharevousa lexicon, which may explain why they are not affected by the same processes as the voiceless clusters.
dominates faithfulness. I will then show that, although ill-formed on the manner dimension, a subset of SF clusters can surface thanks to a constraint that preserves the feature [strident], which, itself, outranks *SO. I then consider the neutralization process affecting the marked obstruent clusters, SF, FF and SS. I argue that the neutralization process to the unmarked FS corresponds to the repair strategy to prevent ill-formed clusters from surfacing in Modern Greek. This is shown to result from the ranking of Type 1 languages.

3.2.4.1 Neutralization of SF Clusters and their Ill-formedness

In this section I argue that, although a restricted subset of SF clusters occurs in dimotiki, these types of clusters, in general, are ill-formed with respect to the constraint system defining obstruent clusters. There are two arguments supporting their ill-formedness. Firstly, unlike FS in which any fricative in the language can precede a stop, e.g. /sp st sk ft xt fk/, in the case of SF clusters only /s/ can follow a stop, e.g. /ps ks/ but */px kf/. Secondly, when SF clusters are created via morpheme concatenation, except when the fricative is /s,z/, they are neutralized to the unmarked FS (Kaisse, 1989). This process is shown in the data below:

(3) /paralei[p+θ]ika/ → [paraleiftika] I was neglected
/kata[ð]io[k+θ]ika/ → [kata[ð]iotika] I was pursued
/ple[k+θ]ika/ → [plextika] I was knitted
The process exemplified in (3) is quite unexpected if SF clusters are indeed considered well-formed in the language. Unlike FF and SS clusters which undergo a similar process that could easily be justified as a dissimilation process, in the case of SF clusters the process in (3) could not be analyzed as dissimilation. The segments affected do not, indeed, share the same value for the feature [continuant]. On the other hand, it could not be explained as a metathesis process due to the fact that the place features are not metathesized either. Such a process would be hard to justify if SF clusters were considered well-formed with respect to the obstruent cluster constraints.

I argue, therefore, that this process, which I characterize as neutralization to the unmarked FS, is evidence that SF clusters are ill-formed in the dimotiki sub-lexicon with respect to the constraints defining obstruent clusters. This process defines the repair strategy for ill-formed clusters. It results from the constraint ranking that makes SF clusters ill-formed with respect to the constraint hierarchy defining obstruent clusters. The relevant ranking is given below:

(4) *SO >> Ident(cont)

Recall that also FF and SS are ill-formed in dimotiki. The fact that FF and SS are ill-formed means that all three markedness constraints dominate Ident(cont) in the dimotiki sub-lexicon. An input containing an SF cluster can, therefore, only surface as an FS cluster regardless of its faithfulness violations. A summary of the ranking is shown in the following tableau:
In the above tableau, the FS cluster best satisfies the constraint system because it incurs two minimal violations of the Ident(cont) constraint. The SF cluster that arises from morpheme concatenation is therefore repaired via simultaneous changing of the feature [continuant] values on both segments in the cluster. The ill-formed SF cluster is hence neutralized to the well-formed FS, i.e. the unmarked obstruent cluster (candidate b).

As shown, the neutralization process that affects SF clusters arising from affixation shows that SF clusters are indeed ill-formed in Modern Greek. If they were well-formed there would be no reason why this process should apply. It must still be kept in mind however, that a restricted subset of SF clusters, i.e. STOP+s clusters, is not only found in monomorphemic words but is also the product of affixation. I argue later in the chapter that this clusters indeed constitute a “protected” subset of SF clusters because of a special faithfulness constraint on the feature [strident].

In the next section, I focus on the dissimilation processes affecting FF and SS clusters that arise in the *dimotiki* sub-lexicon because of morpheme
concatenation or borrowing from the *katharevousa* sub-lexicon. I argue that such dissimilation processes can also be interpreted as neutralization of ill-formed clusters to the unmarked FS. I show that they follow from the same type of interaction that motivates neutralization of SF clusters.

### 3.2.4.2 Neutralization of FF

The first process is a dissimilation process that applies to a sequence of two non-strident fricatives and creates an FS clusters. In this case the second fricative becomes a stop. The process applies to monomorphemic words of Postclassical Greek origin as shown in the data in (6), as well as in morphologically complex words resulting from affixation of a fricative-initial suffix to a fricative-final stem, as shown in the data in (8):

(6) Fricative dissimilation in Postclassical Greek words

<table>
<thead>
<tr>
<th>Postclassical</th>
<th>Dimotiki</th>
</tr>
</thead>
<tbody>
<tr>
<td>sxolio</td>
<td>skolio</td>
</tr>
<tr>
<td>xθes</td>
<td>xtes</td>
</tr>
<tr>
<td>fθinos</td>
<td>ftnos</td>
</tr>
</tbody>
</table>

(7) Fricative dissimilation from morpheme concatenation

/γράφθουκε/ → [γράφτικε]  ”it has been written”

/fιλάχθουκε/ → [fιλάχτικε]  ”it has been honored”
The data in (6) and (7) show that the process applies to both word-initial as well as medial clusters, which, as stated earlier, could either be syllabified as an onset cluster or a coda-onset cluster. However, the fact that both (6) and (7) show the same type of change, suggests also that the clusters in (7) are tautosyllabic onsets.

I will refer to this process as follows:

(8) Neutralization of FF

\[ \text{FF} \rightarrow \text{FS} \]

I argue that this process is motivated by the ranking that defines the ill-formedness of FF clusters, as given below:

(9) \text{OCP}[+\text{cont}] >> \text{Ident(cont)}

This ranking accounts for the neutralization process of FF clusters in monomorphemic and well as morphologically complex words, respectively in tableaux (10) and (11):

(10)

<table>
<thead>
<tr>
<th>/x\theta es/</th>
<th>OCP[+cont]</th>
<th>Ident(cont)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. x\theta es</td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>b. ⠹ ⠙ x\text{tes}</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

In tableau (10), the Postclassical Greek borrowing containing an ill-formed cluster surfaces in Modern Greek as a cluster of a fricative and a stop because of the minimal violation incurred by the unfaithful candidate (b).
In tableau (11), the neutralization process is shown to occur when the offending cluster is formed via morpheme concatenation. The optimal candidate (b) contains an FS cluster. Note also that due to the ranking *SO >> Ident(cont), candidate (c) is not a possible surface form for an input containing an FF cluster.

3.2.4.3 Neutralization of SS

The second process is a dissimilation process that affects sequences of two stops and creates an FS cluster. The process applies to words from Classical Greek origin with such clusters, as well as to clusters arising due to morpheme concatenation as shown in the following examples.

(12) Stop dissimilation in Classical Greek words

<table>
<thead>
<tr>
<th>Classical</th>
<th>Dimotiki</th>
</tr>
</thead>
<tbody>
<tr>
<td>hepta</td>
<td>hefta</td>
</tr>
<tr>
<td>ktizo</td>
<td>xti zo</td>
</tr>
<tr>
<td>okto</td>
<td>ox to</td>
</tr>
<tr>
<td>pteron</td>
<td>ftero</td>
</tr>
</tbody>
</table>
(13) Stop Dissimilation from morpheme concatenation

/plek+ta/ → [plexta] "knitwear"

/lep+ta/ → [lefta] "money"

I will refer to this process as in (14) below:

(14) Neutralization of SS:

SS → FS

The dissimilation process observed with sequences of two stops follows from the same type of interaction between the markedness constraint relevant to SS clusters and the faithfulness constraint. This is shown in the following tableau:

(15)

<table>
<thead>
<tr>
<th>/ktizo /</th>
<th>OCP[-cont]</th>
<th>Ident(cont)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ktizo</td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>b. xτizo</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

In tableau (15) an input containing a cluster of two stops shows neutralization to the unmarked FS due to the fact that the violation incurred by the candidate containing an unfaithful cluster of the type FS is a lesser violation than the one incurred by the faithful candidate containing an SS candidate.

An SS cluster resulting from affixation undergoes the same type of neutralization due to its ill-formedness in the *dimotiki* sub-lexicon, as shown in the tableau below:
The three processes just described produce the same output, i.e. a sequence of the FS type. In linear phonology terms, this is a clear case of conspiracies, i.e. different rules with the same phonotactic function (Kisseberth 1970). In the case of Modern Greek the conspiracy consists in the fact that three separate rules converge to create the same effect, i.e. neutralization of the marked types of obstruent clusters, i.e. FF, SF and SS, to the unmarked type FS. This situation is schematically represented in (17):

(17)  
```
    FF
  SS   FS
  SF
```

The phonological processes observed in the dimotiki lexical stratum strongly suggest that dimotiki is, indeed, a Type 1 grammar, i.e. a grammar in which only FS clusters are well-formed obstruent clusters. The ranking that accounts for the neutralization processes observed is therefore the same as the one defining Type 1 languages:

(18) OCP[+cont], OCP[-cont], *SO >> Faithfulness
This ranking, where all the markedness constraints dominate Faithfulness, ensures that only the unmarked FS clusters will be able to surface due to the fact that such clusters do not violate any of the markedness constraints. However, when ill-formed clusters arise in the language because of borrowings from other sources or morphological processes as in the case of Modern Greek, then the ill-formed clusters are repaired in some way. The repair strategy depends on which faithfulness constraint is lower ranked. In the *dimotiki* grammar, the repair strategy does not delete or add anything in the ill-formed sequences, but rather changes input values of the feature continuant so as to neutralize the ill-formed sequences to FS. The relevant constraint that accounts for the neutralization process is Ident(cont). Interaction of this constraint with the markedness constraints introduced previously provides a unified account of the three main processes affecting obstruent clusters in *dimotiki*.

To conclude, the three processes affecting obstruent clusters in *dimotiki* all follow from the assumption that the language is, in fact, a *Type 1* language. The ranking in which the three markedness constraints dominate faithfulness directly accounts for neutralization of all the marked types of clusters to the unmarked type FS. The constraint system proposed for obstruent clusters allows us to provide a unified account of the three processes just described.
3.2.5 STOP+s Clusters

We need now to explain the behavior of STOP+s clusters. I have shown that SF clusters are ill-formed with respect to the hierarchy defining obstruent clusters and are repaired into the unmarked FS when created by affixation. Monomorphemic words within the *dimotiki* lexicon are not neutralized to FS sequences if the SF cluster consists of a stop followed by /s/. The following words constitute regular *dimotiki* phonotactics:

(19):

- psari "fish"
- tsai "tea"
- ksenos "stranger"

Moreover, STOP+s clusters are even created, if in a sequence of two fricatives, the second one is a strident, i.e. /s/ or /z/. In this case the first fricative becomes a stop and the strident does not change. The process is illustrated in the following examples:

(20)

\[
\begin{align*}
/\text{γράφ}+\text{s}]/ & \rightarrow [\text{γράψ}]/ \quad \text{"I write" (Perfective non-past)} \\
/\text{διάλεγ}+\text{s}]/ & \rightarrow [\text{διάλεκς}]/ \quad \text{"I write" (Perfective non-past)}
\end{align*}
\]

As pointed out by Kaisse (1989), clusters such as the ones in the first two examples arise quite frequently since both future and simple past active morphemes in Modern Greek begin with /s/ and a large number of roots ends in

---

4 In Modern Greek obstruent clusters must agree in voicing. This explains the devoicing of /γ/ in [διάλεκς].
fricatives. The question is then, given the constraint ranking of Modern Greek, how do we account for the occurrence of this particular subset of SF clusters in such a grammar? In other words, we need to explain why an FF cluster in which the second member is a sibilant surfaces as an SF cluster rather than the expected FS cluster. For example, given underlying /γράφ+σο/, we would expect it to surface as [γράφ+το] rather than [γράψο], as shown in tableau (21):

(21)

<table>
<thead>
<tr>
<th></th>
<th>*SO</th>
<th>OCP[+cont]</th>
<th>Ident(cont)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. γράφσο</td>
<td></td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>b. γράφτο</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>c. γράψο</td>
<td>*!</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

Tableau (21) shows that candidate (c), which is the desired output, loses because of its violation of dominant *SO. The system thus would falsely predict that, also in this case, an FS cluster should be the output of the process regardless of the segmental composition of the cluster. In order to prevent candidate (b) from surfacing as the optimal output for an input such as the one in (21), I propose that a correspondence constraint of the Ident(F) family (McCarthy & Prince 1995) is active in the grammar of Modern Greek. This constraint preserves input /s,z/ even if it would result in a violation of the *SO constraint. According to Lombardi (1995), the relevant feature that distinguishes these two sounds from all
other fricatives in Modern Greek is [strident]. I assume that only fricatives are specified for this feature, whereas stops are not (Lombardi 1995). The constraint is formulated according to the general schema of Ident(F) constraints (McCarthy and Prince 1995):

(22) Ident(strident)

Let $\alpha$ be a segment in $S_1$ and $\beta$ be any correspondent of $\alpha$ in $S_2$. If $\alpha$ is $[\gamma]_{\text{strident}}$, then $\beta$ is $[\gamma]_{\text{strident}}$.

I am assuming that this constraint is violated under the following two circumstances:

(a) $[\alpha]_{\text{strident}} \rightarrow [\beta]_{\text{strident}}$

(b) $[\alpha]_{\text{strident}} \rightarrow [\varnothing]_{\text{strident}}$

With this additional constraint we can easily explain the presence of [ps] and [ks]

---

5 Greek has both /θ, δ/ and /s, z/. These sounds are described as apico-interdental and apico-dental. Lombardi (1995) argues that since both sounds are apical, they cannot be distinguished by this property. If these sounds are not distinguished by the feature [strident], then an additional place distinction must be introduced, so that a distinction between interdental and dental would be possible. Moreover, if an apical/laminal distinction at various places of articulation is kept, the system would predict an apico/dental/apico-interdental contrast in the stops. But such a contrast is never observed. Lombardi concludes, therefore, that there is no real reason why fricatives should show more distinctions than stops and consequently the distinguishing feature is [strident].

6 McCarthy and Prince (1995) distinguish between two types of assessment violations for Ident(F). In the case of binary features there is a violation if (a) $+F \rightarrow -F$ or (b) $-F \rightarrow +F$. In the case of privative features violations are assessed in the following two cases: $F \rightarrow \varnothing$ or $\varnothing \rightarrow F$ (see Lombardi 1999 for discussion). Here we are evaluating a binary feature which is only specified for certain segments types. For this reason there is a combination of the two types of evaluation. Note also that the constraint is violated if $\varnothing \rightarrow [\alpha]_{\text{strident}}$. This situation is not relevant for the case at hand.
clusters in Modern Greek regardless of their violation of the *SO constraint. The interaction is shown in the following tableau:

(23)

<table>
<thead>
<tr>
<th>/γράφ+so/</th>
<th>Ident(strident)</th>
<th>OCP[+cont]</th>
<th>*SO</th>
<th>Ident(cont)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. γράφο</td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>b. γράφτο</td>
<td>*!</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>c. γράψο</td>
<td></td>
<td>*</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

Candidate (c) is now the optimal candidate regardless of the fact that it violates *SO. Candidate (a) fails because of its violation of OCP[+continuant], whereas candidate (b), containing an unmarked FS cluster, loses because of the fact that the [+strident] feature associated with the segment /s/ in the input, is not present in either segments of the output. The fricative [f] is a [-strident] segment, whereas [t] is not specified for the feature at all because I am assuming stops are unspecified for stridency. In other words, there is no correspondent of /s/ in candidate (b) that carries the feature [+strident]. Regardless of the fact that this candidate corresponds to an unmarked FS cluster, it is not optimal because of the absence of a [+strident] segment in the output. Notice also that OCP[+cont] must
crucially dominate *SO in order for (c) to be optimal. On the other hand, from this tableau no relative ranking between Ident(strident) and OCP[+cont] can be determined.

I have shown, so far, that Modern dimotiki Greek is a Type 1 language in terms of the typology of obstruent clusters. Moreover, I have argued that one particular type of SF clusters, i.e. STOP+s, are notwithstanding allowed to surface under duress by Ident(strident). The constraint ranking for dimotiki Greek is therefore summarized as follows:

(27) Ident(strident) OCP[-cont] [OCP+cont]

\[ \text{Ident(cont)} \]

\[ \text{*SO} \]

The constraint ranking in (27) accounts only for the phonotactics and phonology of obstruent clusters in the dimotiki lexical stratum. However, nothing has been said, so far, for the clusters allowed in the katharevousa sub-lexicon and how the two lexical-strata can co-exist in Modern Greek as a whole. Lexicon stratification is potentially a problem because the two lexical strata show different phonotactics and phonologies that cannot obviously result from a single constraint ranking. In what follows I will describe how the model of lexicon stratification independently proposed by Fukazawa (1997, 1999), Fukazawa, Kitahara and Ota
(1998) and Itô and Mester (1998) can handle diaglossic languages in a straightforward way.

3.2.6 Lexicon Stratification and Modern Greek

In the vast majority of languages, the lexicon at large can be partitioned into separate sub-lexicons based on etymology. A fairly standard partitioning is the one between the native vocabulary and one or more separate loanword lexicons. These different sub-lexicons usually show different phonotactic restrictions or undergo stratum-specific phonological processes. Fukazawa (1997, 1999), Fukazawa, Kitahara and Ota (1998) and Itô and Mester (1998) independently propose to model lexicon stratification, and hence the stratum-specific phonologies, by arguing that faithfulness constraints can be relativized to each individual stratum, and ranked independently with respect to the various markedness constraints active in a given grammar. Under this model, the co-existing but different phonotactics and phonologies of the katharevousa and dimotiki sub-lexicons of Modern Greek could be explained by relativizing the Ident(cont) constraint to the two different lexical strata of Modern Greek. In other words, the grammar of Modern Greek would consist of two separate Ident(cont) constraints, one applying only to words in the dimotiki sub-lexicon (Ident(cont)dimotiki), and another to words of the katharevousa sub-lexicon (Ident(cont)katharevousa). The relevant ranking of these constraints would be as in (25) below:
However, making use of an indexed faithfulness constraint, i.e. \(\text{Ident}^{\text{dimotiki}}\) to capture the core behavior may miss the important generalization that this constraint is actually the general constraint of the language at large. In other words, in the case at hand, obstruent clusters neutralization in a lexical item from a third source will not be triggered by low ranking \(\text{Ident}^{\text{dimotiki}}\) because such a lexical item is not part of the \(\text{dimotiki}\) lexicon. The situation of languages such as Modern Greek can best be described by assuming that there are two \(\text{Ident}^{\text{cont}}\) constraints. One \(\text{Ident}^{\text{cont}}\) constraint is the general one, which applies throughout the language regardless of etymology. The other \(\text{Ident}^{\text{cont}}\) constraint is, instead, a special constraint that only applies to words of the \(\text{katharevousa}\) sub-lexicon. Under this view then, the ranking that will give rise to the asymmetry between the two lexicons, will follow directly from the fact that the two constraints stand in a special (\(\text{Ident}^{\text{katharevousa}}\)) to general (\(\text{Ident}^{\text{dimotiki}}\)) relationship, and the effects of the special can only be seen when the special dominates the general. With this ranking than any borrowed lexical item with an ill-formed cluster will undergo neutralization to the unmarked FS except words of the \(\text{katharevousa}\) sub-lexicon. In tableau (26), a common colloquial \(\text{katharevousa}\) word containing an SS cluster does not undergo neutralization to the unmarked FS because it is protected by the special \(\text{Ident}^{\text{katharevousa}}\) which dominates the relevant markedness constraints.
Tableau (26) shows that a *katharevousa* lexical item containing an ill-formed cluster in Modern Greek can survive neutralization due to the effect of high ranking $\text{Ident}(\text{cont})^{\text{kath.}}$. In other words, high ranking $\text{Ident}(\text{cont})^{\text{katharevousa}}$ will allow all types of obstruent clusters to surface in Modern Greek, as long as they are part of the *katharevousa* sub-lexicon. Borrowings from other sources will not enjoy the same privilege because they are only subject to the low ranking general $\text{Ident}(\text{cont})$. The constraint ranking for modern Greek can be therefore summarized in (27) below.

(26)

<table>
<thead>
<tr>
<th>/pteriya/ $\text{kath.}$</th>
<th>$\text{Ident}(\text{cont})^{\text{kath.}}$</th>
<th>OCP[-cont]</th>
<th>$\text{Ident}(\text{cont})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. pteriya</td>
<td>*!</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>b. $\in$ pteriya</td>
<td></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

(27)

```
Ident(\text{cont})^{\text{katharevousa}}
```

```
Ident(\text{strident}) \quad \text{OCP[-cont]} \quad [\text{OCP+cont}]
```

```
\ast \text{SO}
```

```
\text{Ident}(\text{cont})
```

To conclude, Modern Greek exemplifies a grammar of *Type 1* in which ill-formed clusters are repaired by neutralization to the unmarked FS. This results
from the interaction of the markedness constraints with the faithfulness constraint Ident(cont). Interestingly, on the surface Modern Greek could be classified as a less restrictive type of language with respect to the obstruent cluster typology. However, a number of phonological processes affecting obstruent clusters in this language lead to the conclusion that Modern Greek is necessarily a Type 1 language. Moreover, the Greek data provides crucial evidence for the activity of *SO, the new constraint proposed in this typology.
3.3 Case Study II: Lushootseed-Nisqually

3.3.1 Introduction

Lushootseed-Nisqually (Hoard 1978) is one of the southern dialects of Lushootseed, a Native American language of the Salish family spoken in the vicinity of Seattle, Washington state. This language represents an interesting system in that it exemplifies a different type of repair strategy for ill-formed clusters than the ones considered in Chapter 2. In particular, in Nisqually ill-formed obstruent clusters are repaired by obstruent syllabicity. I will first discuss onset obstruent clusters and obstruent syllabicity in Lushootseed-Nisqually. I will argue that not only the system I propose can account for the well-formedness of all obstruent clusters except SS clusters in Nisqually, but also that obstruent syllabicity, i.e. the repair strategy to avoid SS clusters to surface, is completely predictable from basic syllable structure constraints and the system of constraints proposed for obstruent clusters.

3.3.2 Obstruent Clusters and Syllabic Obstruents in Nisqually

Nisqually’s obstruent inventory is given in Chart (28)7.

---

7 The chart is based on a limited set of data provided in Hoard (1978). /h/ and /g/ do not occur in the data. I assume that the lack is just an accident of the data, since all the closely related dialects (Snyder 1968; Hess 1977; Urbanczyk 1996) contain such segments. Moreover, the absence of these segments would be odd given that the language contains a velar rounded and glottalized series. The other dialects also contain the voiced affricates /dz dʒ/ and the voiceless lateral affricate /tɬ/. The data provided in Hoard, however, does not contain such segments at all. I will therefore not include them in the chart.
This language contains sequences of up to three adjacent obstruents. Of these, only FS, SF and FF sequences may be tautosyllabic onset clusters.

(29) Examples of well-formed onset obstruent clusters

- **FS:**
  a) \([sk^w]\acute{\text{a}} \ \text{wəl}\) “stealhead”
  b) \([lq']\acute{\text{a}}\text{či}\) “one-armed man”

- **SF:**
  c) \([t\chi^w]\acute{\text{a}}\text{q}^w\) “Mt. Rainier”
  d) \([q^w\chi^w]\acute{\text{a}}\text{či}\) “fingernail”
  e) \(\text{sčát}\.[q]\text{či}\) “mountain lion”

- **FF:**
  f) \([s\chi]\acute{\text{a}}\text{či}č\) “sword fern”

---

8 Note that the data provided in Hoard does not contain sequences consisting of either the velar or uvular fricatives as the leftmost element in the clusters. However, there is no reason to believe that such clusters are disallowed given the quite unrestricted clustering possibility. In addition, such clusters do occur in the northern dialects.
In contrast, sequences of two stops are never parsed into the same syllable. These sequences are always parsed into two separate syllables. One of the stops is parsed as a syllabic stop and the other will be parsed as either the onset of the following syllable (30a,b) or the coda of the preceding one (30c).

(30) Examples of syllabification of sequences of two stops

a) traî.qá.čiʔ    “eight”
b) ɛ’.bálʔ.qid    “mink”
c) ḥóq’.t        “wide”

Note that affricates pattern with stops since they never occur in the same syllable with another stop. This follows straightforwardly from the treatment of affricates proposed in Lombardi (1990) as shown in section 3.3.1.2 below.

Sequences of three obstruents never form tautosyllabic clusters regardless of whether they contain adjacent stops or not. These sequences are also split into separate syllables. When in word initial or final position, these sequences form heterosyllabic syllables with either syllabic fricatives or fully released syllabic stops as syllable nuclei. This is shown in (31) below:
Examples of syllabification of sequences of three obstruents

a) č’êč’.qs “mosquito”
b) txʷ.č’îχ “stingy”
c) txʷ.só. bəd “bee”
d) sqʷół.ps “cutthroat trout”
e) t’áq.t’.qəc “vine maple”
f) tágʷ.txʷ “moon”
g) ?u.č’i.lə.pal.bšt “He saddled his horse”

The data shown in (29), (30) and (31) lead to the conclusion that within the Typology of Onset Obstruent Clusters, Nisqually is a Type 4 language. In Type 4 languages, FS, SF, FF sequences are well-formed surface onset clusters, whereas SS are ill-formed and cannot surface as tautosyllabic clusters. I propose that the syllabification patterns observed in Nisqually are forced by the constraint ranking established for Type 4 languages. Obstruent syllabicity is the repair strategy adopted to prevent ill-formed clusters from surfacing as tautosyllabic onsets. This results from the interaction of the constraints on obstruent clusters and the Peak Hierarchy (Prince and Smolensky 1993). This hierarchy, together with the Margin Hierarchy, has been proposed to account for what segments can be possible peaks or onset/codas in a language.
3.3.3 The Analysis

3.3.3.1 2-Obstruent Sequences

Based on the data provided in the previous section, Nisqually is a Type 4 language. Grammars for languages like Nisqually, that allow tautosyllabic sequences of FS, SF and FF but disallow SS clusters, have the following constraint ranking:

(32) OCP[-cont] >> Faith >> OCP[+cont], *SO

This constraint ranking assures that FS, SF and FF clusters are well-formed, thus are able to surface tautosyllabically. This same ranking also assures that SS clusters are ill-formed outputs and can never surface as tautosyllabic. Nisqually repairs sequences that would lead to violations of high ranked OCP[-cont] by syllabifying the two stops into different syllables. If these sequences occur either word initially or finally, one of the stops will acquire syllabic identity and constitute a syllable peak.

Prince and Smolensky (1993), in their Typology of Onset and Nucleus Inventories, distinguish between willing and coercible peaks. Willing peaks are peak-preferring tenable peak segments, whereas coercible peaks are those margin-preferring segments that can be forced to be parsed as peaks by higher ranking syllable structure constraints. Under the Containment Model (Prince and Smolensky 1993), the necessary and sufficient condition for a segment $\alpha$ to be a coercible peak is that

(33) PARSE, FILL$^\text{Nuc}$ >> *P/$\alpha$ >> *M/$\alpha$
Ranking PARSE above *M/α ensures that α be a possible surface onset segment in the language. The fact that α is a margin-preferring segment is captured by the domination relation *P/α >> *M/α. For α to be forced to surface as a peak, it is necessary that a syllable structure constraint, in this case FILL Nuc, dominate the anti-association constraints.

Following this line of reasoning, Nisqually’s obstruents are coercible peaks. They are margin-preferring segments coerced in the peak by the higher ranking constraint OCP[-cont].

Under Correspondence Theory (McCarthy and Prince 1995), Condition (33) can be reinterpreted as (34) to apply to Nisqually.

(34) MAX-IO, DEP-IO >> *P/Ob >> *M/Ob

In (34) MAX-IO and DEP-IO take the place of PARSE and FILL Nuc respectively. DEP-IO should be understood as actually DEP-IO Nuc. As in the Parse/Fill model, distinguishing between DEP-IO Nuc and DEP-IO Ons seems necessary in order to be able to derive the “Extended CV Syllable Structure Typology” (see Prince and Smolensky (1993), chapter 6). *P/Ob and *M/Ob are composite constraints that refer to the natural class of obstruents. These constraints encapsulate the sets of anti-association constraints for the members

---

9 *M/Ob will not be discussed here since it is ranked at the bottom of the hierarchy and is always violated.
of the class of obstruents\textsuperscript{10}. In particular, \( ^*P/\text{Ob} \) abbreviates the universal ranking \( ^*P/\text{Stop} >> ^*P/\text{Fricative} \) proposed in Prince and Smolensky.

In the case of Nisqually, obstruent syllabicity is forced by the syllable structure constraint OCP[-cont], which must dominate the anti-association constraint against parsing an obstruent as a peak. The ranking OCP[-cont] >> \( ^*P/\text{Ob} \) will satisfy the necessary condition for obstruent syllabicity to be possible. The interaction between the two constraints is shown in the following tableau.

\begin{table}
\begin{tabular}{|c|c|c|}
\hline
\text{Candidate} & \text{OCP[-cont]} & \text{\( ^*P/\text{Ob} \)} \\
\hline
a. \text{.tqá.čiʔ} & *! & \\
\hline
b. \text{.tq.á.čiʔ} & *! & * \\
\hline
c. \text{.tqá.čiʔ} & * & * \\
\hline
\end{tabular}
\end{table}

Candidates (a) and (b) violate OCP[-cont] since in both forms the two input stops are tautosyllabic. The only difference between the two candidates being that in candidate (a) the two obstruents are in the onset, whereas in candidate (b) the first obstruent is parsed as the onset of a syllable whose peak is a syllabic stop. In contrast, the winning candidate avoids violation of the syllable structure.

\textsuperscript{10}In Prince and Smolensky (1993), it is assumed that the two classes of segments have different sonority values, with fricatives being more sonorous than stops. Based on sonority, they propose universal rankings of anti-association constraints, that can be encapsulated as \( ^*P/\text{Stop} >> ^*P/\text{Fricative} \) to account for core syllabification in Imdlawn Tashlhiyt Berber (Dell and Elmedlaoui 1985, 1988, 1989; Prince and Smolensky 1993; Clements 1997). This view can be reconciled with the present work by assuming that some intrinsic property of these segments other than "sonority" may be responsible for making fricatives better peaks that stops. Determining what this property may be is, however, outside the scope of the dissertation.
constraints against tautosyllabic stops by breaking the sequence into two different syllables, the first of which consists of a syllabic stop. Recall from Chapter 2, that the three markedness constraints, i.e. OCP[-cont], OCP[+cont] and *SO, range over tautosyllabic but not heterosyllabic clusters.

Given that Nisqually is a Type 4 grammar, OCP[-cont] must be ranked higher than all the faithfulness constraints (Indent(cont), MAX-IO and DEP-IO) to disallow sequences of two stops to surface faithfully as tautosyllabic clusters. The faithfulness constraints are unranked with respect to each other since neither deletion, epenthesis or feature specification changing is adopted as a repair strategy to resolve inputs that would lead to ill-formed outputs. Moreover condition (34) requires that MAX-IO and DEP-IO be ranked above the anti-association constraints, therefore the following constraint ranking can be established:

(36)  \( \text{OCP[-cont]} \gg \text{Indent(cont)}, \text{MAX-IO, DEP-IO} \gg \text{*P/Ob} \gg \text{*M/Ob} \)

Tableau (37) shows the interaction of the faithfulness constraints with *P/Ob.

(37)

<table>
<thead>
<tr>
<th>/tqa ~/</th>
<th>Ident(cont)</th>
<th>MAX</th>
<th>DEP</th>
<th>*P/Ob</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ʕʔ qa ~</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>b. qa ~</td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>c. tə qa ~</td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>d. sqa ~</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Any violation of the faithfulness constraints is worse than parsing one of the stops as a peak. Obstruent syllabicity is therefore less costly than either deletion (candidate b) or insertion (candidate c) of a segment to break up the input sequence; or changing the feature specifications of the segments and turning the input sequence into the unmarked cluster type, i.e. a FS cluster (candidate d).

The constraints that penalize clusters of the form FF and SF, i.e. OCP[+cont] and *SO must be dominated by the anti-association constraint *P/Ob to assure that these sequences can surface as tautosyllabic clusters rather than being split into separate syllables with syllabic obstruents. The following two tableaux show an input containing a sequence of two fricatives and one containing a stop and a fricative respectively.

Tableau (38) shows that an input sequence of two fricatives will surface as an onset cluster. This is possible because the ban against two tautosyllabic fricatives is outranked by the anti-association constraint that disallows to parse an obstruent as a peak. It is therefore better to parse the sequence as an onset cluster rather than splitting it in two separate syllables.
In (39) an input sequence of the form SF is shown. This case also shows that a violation incurred by parsing one of the obstruents as a peak is worse than parsing the sequence as a tautosyllabic cluster.

### 3.3.3.2 Sequences Containing Affricates

Affricates\(^\text{11}\) are allowed to form clusters with fricatives but not with other stops, as shown in the following examples:

\[(40)\]

(a) \([\text{ʃ}\text{ʃ}]\text{āb.\k}\) “cloud”

(b) \(\text{č’}.\text{bál?}.\text{qid}\) “mink”

This restriction follows from Lombardi’s ‘unordered affricate’ (1990). Lombardi proposes that affricates are composed of [-cont] and [+cont] specifications which are unordered at underlying representation and throughout the phonological derivation, although they are ordered phonetically. She argues that these

\(^{11}\text{If the voiced affricates mentioned in footnote 3 are part of the phonemic inventory of the language, the rest of the data provided suggests that they should occur in clusters. This expectation is based on the observation that voiceless affricates freely combine with other obstruents and that Nisqually, unlike the dialect described in Urbanczyk (1996), seems to allow clusters with different voicing specifications.}\)
segments are represented with the different values of [cont] on separate tiers, as in (41): \[12\]

```
(41)      [+cont]  
          |     Root
          |    [-cont]
       Root
```

as opposed to the representation in (42) in which both specifications are on the same tier (Sagey 1986 among others):

```
(42)  X  
      |  Root
      |  \
    Root  /  \  [-cont]  [+cont]
```

The representation in (41), unlike the one in (42), correctly predicts that a sequence that contains an affricate and a stop (40b) is syllabified as an heterosyllabic sequence. OCP[-cont] prohibits a tautosyllabic sequence of two [-cont] specifications. Only under Lombardi’s proposal the two specifications for [-cont] in a sequence containing an affricate followed by a stop would be adjacent and hence subject to OCP[-cont] as illustrated below:

```
(43)      [+cont]  
         |   c'  b ~  
         |     |  
[-cont] [-cont] \   [OCP[-cont]]
```

---

12 Lombardi (1990) concludes that [+cont] and [-cont] are actually two separate privative features [continuant] and [stop]
If, on the contrary, the [-cont] and [+cont] specifications for an affricate are on the same tier, then the two [-cont] specifications of the sequence would not be adjacent and there would be no OCP violation. We would incorrectly predict such a sequence to surface as a tautosyllabic cluster. This is illustrated in (44) below:

\[
\begin{array}{ccc}
  & c' & b \\
/ & \backslash & ~ \\
\hline
[-cont] & [+cont] & [-cont]
\end{array}
\]

Under Lombardi’s proposal, a sequence containing a fricative and an affricate (therefore two adjacent [+cont] specifications), will still be able to surface tautosyllabically, as in the case of any sequence of two fricatives, because of the ranking *P/Ob >> OCP[+cont]. This ranking allows [+cont] sequences to form tautosyllabic clusters rather than requiring a syllable break between the two.

### 3.3.3.3 3-Obstruent sequences

Clusters consisting of three consonants are never allowed, regardless of whether they consist of three obstruents or of segments with decreasing sonority values. I therefore assume an undominated constraint that bans three consonant clusters all together. The constraint is formulated as in (45):

\[
*3-C \text{ (Three consonant clusters are disallowed)}^{13}
\]

---

\(^{13}\)The nature of this constraint is quite unclear at the moment.
As a consequence of such a more general restriction on complex syllables, three obstruent clusters are ill-formed regardless of their form. When the sequence, however, contains sub-sequences that form well formed clusters themselves, the question arises of how they are broken up into smaller tautosyllabic syllables. If the input sequence contains a sub-sequence with two adjacent stops, SSF for example, then high ranking OCP[-cont] will force a syllable break between the two stops, S.SF. An example of a medial sequence containing three obstruents is given in tableau (46).

(46)  
<table>
<thead>
<tr>
<th>/čátqɬəb/</th>
<th>OCP[-cont]</th>
<th>*P/Ob</th>
<th>*SO</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ɬə čát.qɬəb</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>b. čátq.ɬəb</td>
<td>*!</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>c. ɬə čát.q.ɬəb</td>
<td></td>
<td>*!</td>
<td></td>
</tr>
</tbody>
</table>

Tableau (46) shows that, in the case of a medial sequence, the system will select as the optimal candidate the candidate with no violation of OCP[-cont] and no syllabic obstruents, given the fact that both the constraint that penalizes obstruent syllabiccity and the one that penalizes tautosyllabic sequences of two stops are ranked above *SO.

If, however, both sub-sequences form well-formed clusters, then there will be a question as to how three obstruent sequences are syllabified. In particular, given a sequence of the form SFF, for example, in which both sub-sequences SF and FF form well-formed clusters, what determines the optimal
output form? As an example consider an input sequence such as the one given in tableau (47):

(47)

<table>
<thead>
<tr>
<th>/tx⁸sə~/</th>
<th>*P/Ob</th>
<th>OCP[+cont]</th>
<th>*SO</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. φ tx⁸sə~</td>
<td>*</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>b. ť.x⁸sə~</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

The input in tableau (47), can be syllabified as either candidate (a) or candidate (b). Given the constraint ranking established so far and assuming the encapsulated *P/Ob, both candidates are assigned equal marks. Candidate (a), which is the actual output form (backwards hand) and hence should win, is no different in terms of better or worst constraint violations from candidate (b), which should lose in the competition¹⁴. The violations on OCP[+cont] and *SO are equal because the two constraints are not ranked with respect to each other. As a matter of fact, there is no ranking argument for these two constraints in the language as a whole. However, *P/Ob is a cover constraint for the fixed hierarchy "*P/Stop >> *P/Fricative". Given that these two ranked constraints dominate OCP[+cont] and *SO in the constraint system, evaluation of the candidate set by the higher ranked anti-association constraints will always take precedence over the lower

¹⁴ Candidate (b), unlike candidate (a), also violates the syllable structure constraint Onset. The interaction of syllable structure constraints will be shown later in the analysis.
ranked structural constraints.\textsuperscript{15}

The problem in tableau (47) is actually not a problem once *P/Ob is substituted with the sub-hierarchy *P/Stop >> *P/Fricative, as shown in (48):

\textbf{(48)}

\begin{center}
\begin{tabular}{|l|c|c|}
\hline
\textipa{/tx^w\text{\textipa{sə~}}} & *P/Stop & *P/Fricative \\
\hline
a.  \textipa{ʦ\textipa{ŋ\textipa{t}}\textipa{x^w}\text{\textipa{sə~}}} & & * \\
\hline
b.  \textipa{t\textipa{ŋ\textipa{x^w}sə~}} & *! & \\
\hline
\end{tabular}
\end{center}

Candidate (a) now turns out to be the winner, given the fact that it contains a syllabic fricative rather than a syllabic stop, thus better satisfying the Peak Hierarchy. However, obstruent syllabicity in Nisqually, cannot always be explained in terms of better satisfaction of the Peak Hierarchy. As a matter of fact, not all inputs containing a stop and a fricative will surface with a syllabic fricative rather than a syllabic stop. This situation is exemplified in tableau (49), where an input sequence containing a fricative and two stops surfaces with a syllabic stop rather than a syllabic fricative.

\textsuperscript{15} Bruce Morén and Edward Keer suggested that ranking OCP[+cont] above *SO would pick the right candidate in tableau (47). However, such a solution would fail to account for the data in tableau (48) on the following page because OCP[+cont] and *SO would apply vacuously to both candidates.

\textbf{(48')}

\begin{center}
\begin{tabular}{|l|c|c|c|c|}
\hline
\textipa{/scqā~} & *P/Stop & *P/Fricative & OCP[+cont] & *SO \\
\hline
a. wrong winner  \textipa{ʃc.qā~} & & & * & \\
\hline
b. desired winner  \textipa{ʃc.qā~} & & *! & & \\
\hline
\end{tabular}
\end{center}
In (49), the fixed constraint ranking of the Peak Hierarchy will pick the wrong optimal winning candidate. The candidate that contains a syllabic fricative incorrectly wins over the one that contains a syllabic stop (an affricate in this case). Candidate (a), however, is more marked in terms of syllable structure than (b). Candidate (a) has a coda consonant but lacks an onset, thus violating both NOCODA and ONSET, whereas candidate (b) has an onset but lacks a coda, thus satisfying both syllable structure constraints. This situation is shown in tableau (50).

Tableau (50) shows that either ONSET or NOCODA must crucially dominate the Peak Hierarchy in order to prevent candidate (50a) to win. The question, however, is which of these two constraints is the active constraint in this portion of Nisqually grammar? I propose that ONSET, rather than NOCODA, is active in determining the locus of syllabicity in obstruent sequences. As a matter of fact,
NOCODA in Nisqually must necessarily be ranked low or at least lower than the
Peak Hierarchy given the fact that the language allows simple as well as complex
obstruent codas. Tableau (51) shows an input form in which an obstruent
sequence is syllabified as a coda cluster.

(51)

<table>
<thead>
<tr>
<th>/qədxʷ/</th>
<th>ONSET</th>
<th>*P/Fricative</th>
<th>NOCODA</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. qə.dxʷ</td>
<td></td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>b. qədxʷ</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>c. qəḍ</td>
<td>xʷ</td>
<td>*!</td>
<td></td>
</tr>
</tbody>
</table>

Candidate (51b) wins because it only violates low ranked NOCODA, as opposed
to candidates (54a) and (54c) that do not incur any violations of NOCODA but
lose on the higher ranked constraints. Note also that if NOCODA were ranked
above *P/Fricative candidate (54a) would incorrectly win.

The ranking in which ONSET dominates the Peak Hierarchy and
NOCODA is dominated by the Peak Hierarchy, while completely disactivating
NOCODA, also makes ONSET responsible for determining the locus of
syllabicity in ill-formed obstruent clusters. Given this ranking, the Peak Hierarchy
is also inactive in determining which one of the obstruents will acquire
syllabicity. Another example of ONSET’s activity in determining the locus of
syllabicity is given in tableau (52).
In (52), both candidates (a) and (c) lack an onset, thus both incurring fatal violations of ONSET. Therefore, the winning candidate is the candidate in which the second fricative acquires syllabicity in order to satisfy some higher ranked constraint, in this case undominated *3-C.

The constraint ranking\(^{16}\) for Nisqually is therefore as follows:

\[
\begin{array}{c|c|c}
&s/xlá/ & \text{ONSET} & *P/\text{Fricative} \\
\hline
\text{a. } & \text{š.xlá} & *! & * \\
\text{b. } & \text{sš.lá} & & * \\
\text{c. } & \text{šx.lá} & *! & *
\end{array}
\]

\(^{16}\) The ranking between the higher ranked structural constraints and the faithfulness constraints follows from the typology, although it cannot be established in this particular case. The data provided does not allow to provide a ranking argument for ONSET with respect to the faithfulness constraint.
This constraint hierarchy predicts that onsets in Nisqually will contain maximally two consonants and they will be of the form FS, SF and FF. Sequences of two SS, as well as input sequences containing three consonants, will never be able to surface as tautosyllabic clusters given the constraint ranking. ONSET ranked above the Peak Hierarchy makes obstruent syllabicity completely predictable.

To conclude I have shown that cluster well-formedness in Nisqually can be explained in light of the Typology proposed in this dissertation. The constraint ranking established cross-linguistically directly accounts for the well-formed as well as ill-formed clusters of the language in syllable onsets. Interaction of this hierarchy and the Peak Hierarchy accounts for obstruent syllabicity in the environments where it is observed. Finally, ONSET is argued to be the sole force in determining the locus of syllabicity in sequences that would lead to ill-formed outputs.
CHAPTER 4

PLACE GENERALIZATIONS

4.1 Introduction

The purpose of this chapter is twofold. Firstly, it further contributes to the understanding of phonotactic patterns by presenting place restrictions for obstruent clusters occurring in onset position. Across languages there is a strong preference for clusters consisting of a coronal fricative followed by a stop. I argue that the pattern observed can best be understood as place neutralization in a position of weak perceptibility. This chapter provides, therefore, further support to the idea that phonotactic patterns can be best understood and explained by reference to phonetic facts. The analysis I present draws on work by Beckman (1997), Padgett (1997) and Steriade (1993, 1994, 1995, 1997).

To support the model of grammar that I propose, I will present data from English, German, Delaware and Takelama. English and German represent two standard unmarked cases, in which only coronal fricatives are allowed in pre-obstruent position. The pattern shown by Delaware onset obstruent clusters is particularly interesting in illuminating the phonotactic patterns, since it shows a crucial asymmetry among the obstruents in the same phonological context. In particular, whereas fricatives are restricted to coronal place in pre-obstruent position, stops allow any place of articulation in the same position. Finally, Takelma is an example of a harmonically incomplete language. The two
markedness dimensions relevant to obstruent clusters interact in such a way that more marked obstruent clusters surface at the expense of less marked clusters. In particular the requirement that both stops and fricatives in pre-obstruent position are coronal prevents SF clusters from surfacing, but not SS.

4.2 Onset Place Restrictions

In onset position, I have observed systematic place restrictions mostly for fricatives in pre-obstruent position. The restrictions are listed below:

(a) only /s/ is allowed as first member of the cluster.

(b) any coronal fricative is allowed as first member of the cluster.

(c) any fricative in the language is allowed as first member of the cluster.

This pattern holds regardless of the other types of obstruent clusters that occur in the language. In other words, there are languages, such as Modern Hebrew and Delaware, in which place restrictions hold for fricatives but not stops in the same phonological context. Whereas a fricative in pre-obstruent position, in these languages, must necessarily be a coronal, a stop in pre-obstruent position can have any place of articulation allowed in the language.

To account for the systematic restrictions on first member fricatives and the asymmetric behavior of stops in the same position, I propose a system of constraints which incorporates the notion of segmental release, following on work in the OT literature by Lombardi (1995), Padgett (1997) and Steriade (1997). The system is modeled upon Positional Faithfulness (Beckman 1998). Within this
framework, I show that the place restrictions that hold in onset are indeed the result of place neutralization in a position in which place contrast would be hard to perceive due to the weakness of perceptual cues.

4.3 Release-sensitive Faithfulness and Place Neutralization

Segmental release, i.e. the burst that accompanies the offset phase of a consonantal constriction, has been known to provide important acoustic cues to place contrast and laryngeal features (Ohala (1990); Kingston (1990); Steriade (1993a-b), (1994), (1995a-b); Lombardi (1995); Padgett (1997)). As pointed out by Padgett (1997), release is “virtually phonetically inevitable” in presonorant position, whereas it can be masked by the presence of a following obstruent. This latter position is the environment in which consonants are less likely to be perceived because the perceptual cues are impoverished. Padgett (1997) proposes to implement the idea that features under release are perceptually more salient by release-sensitive faithfulness. His system for place features assumes that all consonants are released before a tautosyllabic sonorant, otherwise they are unreleased. The system consists of a special release-sensitive faithfulness constraint, FAITH$_{REL}$, and a general faithfulness constraint, FAITH$^1$, with the fixed ranking in (1) below:

(1) FAITH$_{REL}$ $>>$ FAITH is universally fixed.

$^1$ In particular Padgett proposes release-sensitive Max and Dep constraints. I will use Ident constraints to be consistent with the other faithfulness constraints in the dissertation.
Following Padgett (1997), I propose that the pattern observed in onset obstruent clusters can be modeled on the basis of the constraints that follow:

(2) **Ident\text{RelPlace}** : Released consonants and their input correspondents must agree in place feature.

This constraint requires that obstruents occurring in a release position, i.e. pre-sonorant, must maintain their input place of articulation. This constraint basically reflects the fact that segmental release contains the strongest cues and is therefore given a prominent status in the system.

The other constraint in the system is a general constraint that ranges over obstruents only, following on Padgett (1997). This constraint is given below:

(3) **Ident-Obstruent-Place**

Obstruent and their input correspondents must agree in place features.

The ranking between these two constraints is universally fixed on the basis of Padgett’s model:

(4)  \text{Ident}_{\text{RelPlace}} \gg \text{Ident-Obstruent-Place}

Finally, I propose that the system contains another special faithfulness constraint, which ranges over stops only and is not universally ranked. This constraint is given in (5) below:

(5) **Ident-Stop-Place**

Stops must agree in place features with their output correspondent
The purpose of this constraint is to preserve input place features for stops regardless of the position in which they occur, whether in a release position or not. This constraint is crucial in the analysis of languages such as Modern Hebrew and Dakota, in which stops, but not fricatives, maintain place contrast in pre-obstruent position. The idea behind this constraint is that internal cues to obstruent place of articulation are stronger for stops than fricatives, unless the latter are sibilants (Wright 1996).

The strength of the stops’ perceptual cues can be captured in the system by the existence of a special constraint that is relativized to stops only. On the contrary, the relative weakness of the fricatives’ perceptual cues is captured by the non-existence of a special constraint, but rather in the fact that fricatives are subject to the general Ident-Obstruent-Place constraint.

From a purely phonological point of view, this constraint finds motivation in inventory considerations. Languages tend to maintain place contrast mostly among the stop series and to a lesser extent in their fricative series. Many languages restrict their fricative series to coronal fricatives only, but they never restrict their stop series to coronal place only. Moreover, there are no languages in which fricatives occur at different points of articulation, but there are only coronal stops in the system. In general the places of articulation for fricatives are a subset of the places of articulation at which stops occur in a language (Maddieson, 1984).
The factorial typology generated by the interaction of these constraints and the constraints that refer to each individual place feature (which I will indicate as \(*F\) for the moment) gives rise to three basic types of languages. Assuming the fixed ranking of (4), we get the following factorial typology:

(6) a. \(*F >> Ident_{RelPlace} >> Ident-Obstruent-Place >> Ident-Stop-Place\)

b. \(Ident_{RelPlace} >> Ident-Obstruent-Place >> Ident-Stop-Place >> *F\)

c. \(Ident_{RelPlace} >> Ident-Obstruent-Place >> *F >> Ident-Stop-Place\)

d. \(Ident_{RelPlace} >> *F >> Ident-Obstruent-Place >> Ident-Stop-Place\)

e. \(Ident_{RelPlace} >> *F >> Ident-Stop-Place >> Ident-Obstruent-Place\)

f. \(Ident_{RelPlace} >> Ident-Stop-Place >> *F >> Ident-Obstruent-Place\)

The ranking in (6a) is the least interesting because it corresponds to a language in which neither fricatives nor stops have the feature \(F\), whether in a release position or not. The rankings in (6b) and (c) characterize a language in which both fricatives and stops occurring in pre-obstruent position allow the feature \(F\). The fact that both rankings give rise to the same type of language is because in both cases \(Ident-Obstruent-Place\), the general constraint, dominates the special \(Ident-Stop-Place\). The rankings in (6d) and (e), instead, characterize a grammar in which neither a fricative nor a stop in pre-obstruent position maintains place contrast. In other words, this might correspond to a language in which both fricatives and stops occurring in pre-obstruent position only allow coronal place. This is actually the case of Takelma, which I will discuss later in the chapter.
Finally, the ranking in (6f) characterizes a language in which only stops but not fricatives can maintain place contrast in pre-obstruent position. This is the case of Delaware and Modern Hebrew.

In the remainder of this chapter I will first provide an analysis of the English and German obstruent cluster phonotactics and show how the restrictions observed for onset clusters can be accounted for with the constraints proposed here. Secondly, I discuss the phonotactics of Delaware, which provides crucial evidence for the proposal that stops are privileged segments over fricatives. I then discuss Takelma which exemplifies both a language in which stops and fricatives in pre-obstruent position are neutralized to coronal, as well as an example of a harmonically incomplete language. Finally, I discuss an alternative analysis based on a strict implementation of Steriade’s Licensing-by-Cue model and show why, with place features, such a system proves to be problematic.

4.4 Case Study III: English

In terms of obstruent clusters, English represents what we could call the unmarked type of language. Obstruent clusters in English are limited to $s+$STOP. These clusters, within the present proposal, represent the best formed types of clusters because they are FS clusters, i.e. the unmarked type on the continuancy dimension. Moreover, on the place dimension, the only fricative that occurs in pre-obstruent position is $/s/$, which is the least marked coronal fricative.
4.4.1 The English Obstruent System

The English obstruent inventory is given in the chart below (O’Grady, Dobrovolsky and Aronoff 1997)

(7)

<table>
<thead>
<tr>
<th></th>
<th>Labial</th>
<th>Interdent.</th>
<th>Alveolar</th>
<th>Alveo-palatal</th>
<th>Velar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stops</strong></td>
<td>p</td>
<td>b</td>
<td>t</td>
<td>d</td>
<td>k</td>
<td>g</td>
</tr>
<tr>
<td>Simple</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Fricatives</strong></td>
<td>f</td>
<td>v</td>
<td>θ</td>
<td>δ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-sibilant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sibilant</td>
<td>s</td>
<td>z</td>
<td>s</td>
<td>z</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Affricates</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-sibilant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sibilant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>tf</td>
<td>dz</td>
</tr>
</tbody>
</table>

In English, the only obstruent clusters that are representative of the native phonotactics consist of the form s+Stop, as the following examples show:

(8) [spɪl] spill
    [stɪk] stick
    [skʌɪ] sky

English is an example of a *Type 1* language in onset, i.e. a language that only admits FS clusters. On the continuancy dimension, the language is characterized by the following ranking:

(9) OCP[-cont], OCP[+cont], *SO >> Ident(continuant)

---

2 Other clusters are also found in English, e.g. in sphere. These clusters are however only found in loans and are not considered representative of English phonotactics.
According to the ranking in (9) only FS clusters are allowed to surface. This was demonstrated in Section 2.2.6 for a Type 1 language.

On the place dimension, it has been observed that only /s/ is allowed in pre-obstruent position. No other fricative occurring in the language is allowed in that position. I have argued in the previous section that this pattern is a pattern of place neutralization in a weak position for place contrast. I have argued that this type of neutralization results from the interaction of the release-sensitive constraint, IdentRelPlace, the general constraint for obstruents, Ident-Obstruent-Place, and the place hierarchy of Prince and Smolensky (1993). The ranking that gives rise to the pattern observed for English obstruent clusters is given below:

(10) IdentRelPlace >> *Dor, *Lab>> Ident-Obstruent-Place , *Cor

Based on the constraint ranking in (10), a potential input of the form /ft/, cannot surface faithfully. On the contrary, the segment [f] will surface as a fricative at the least marked place of articulation, i.e. coronal. We can say therefore that a fricative in pre-obstruent position undergoes place neutralization. The tableau that follows, however, shows that the place hierarchy cannot chose between the three coronal fricatives [s], [ʃ] and [θ] present in the English inventory. This is shown

---

3 Note that since *Cor is the least marked place in Prince and Smolensky, whether it is ranked above or below Ident-Place it will always surface. If, however, we adopt Lombardi’s hierarchy (1995), in which she proposes that *Phar is actually the least marked place, then we would expect [h] to be the preferred sound in that position. However, [h] would be poorly perceptible in that position since it lacks the frication noise typical of fricatives. So I assume that [ht] clusters are ruled out by a higher ranked constraint.
in the tableau by the fact that it contains three optimal candidates.

(11) Tableau illustrating the constraint ranking in (10)

<table>
<thead>
<tr>
<th>/fpV/</th>
<th>Ident_{Rel}(Place)</th>
<th>*Lab</th>
<th>Ident-Obs-Place</th>
<th>*Cor</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>st</td>
<td>*!</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>b.</td>
<td>fp</td>
<td>**!</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td>sp</td>
<td>*</td>
<td>*</td>
<td>**</td>
</tr>
<tr>
<td>d.</td>
<td>fp</td>
<td>*</td>
<td>*</td>
<td>**</td>
</tr>
<tr>
<td>e.</td>
<td>sp</td>
<td>*</td>
<td>*</td>
<td>**</td>
</tr>
</tbody>
</table>

In the tableau, Ident_{Rel}Place ensures that the /p/ maintains its place of articulation when it occurs in a release position (candidate a). The cluster [fp] fatally violates *Lab because it contains a labial in pre-obstruent position and thus loses in the competition. Candidates (c), (d) and (f) are all optimal because they incur identical violations. In the case of English, however, candidate (c) should indeed be the optimal candidate.

To differentiate between these three coronal fricatives, and among the coronal fricatives in general, I propose that *Cor is indeed an encapsulated
constraint which stands for the two independent sub-hierarchies in (12)⁴:

(12)  
   a.  *[-anterior] >> * [+anterior]⁵
   b.  *+[distributed] >> *[-distributed]

The rankings in (11) are based on the assumption that /s/ is, indeed, the least marked of the coronal fricatives. As an example, the fact that [s] is less marked than [ʃ] is captured by the ranking in (12) and shown in the following tableau:

(13)

<table>
<thead>
<tr>
<th></th>
<th>*+[dist]</th>
<th>*[-anterior]</th>
<th>*[-dist]</th>
<th>* [+anterior]</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. s</td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>b. ʃ</td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

The fricative [s] is shown to be less marked than [ʃ] because the latter violates higher ranked constraints. By the same token, the ranking in (12b) shows that

---

⁴ It has been suggested to me that *Cor and the sub-hierarchies in (12) may indeed be independent constraints. Their independence of *Cor may allow them to dominate *Cor. If other constraints intervene, however, this hypothesis can give rise to a ranking that disallowed Coronal place in a language and allowed other places of articulation as below:

(i)  *[-ant], *[+dist] >> *[+ant], [-dist] >> IdentPlace >> *Lab >> *Cor.

To the best of my knowledge, there are no languages that allow labial segments (or dorsal) but not coronal. This therefore discards the possibility that they are indeed separate constraints independent of *Cor.

⁵ These are the features that depend on the Coronal node in Feature Geometry. A problem remains, however, with the typological predictions that follow from these two feature hierarchies. Whether they are ranked with respect to each other as in (a) or (b), or unranked as in (c), the prediction is that across languages the occurrence of [ʃ] implies the occurrence of both [s] and [θ]. This is not, however, an attested pattern because [θ] is more rare than [ʃ]. This prediction may represent a problem for segment typology. For the moment, I assume the ranking in (c).

(a)  *[-ant] >> *[+ant] >> *[+dist] >> *[-dist]
(b)  *[+dist] >> *[-dist] >> *[-ant] >> *+[ant]
(c)  *[-ant], *[+dist] >> *[-dist], *[+ant]
[s] is less marked than the non-distributed coronal fricative, i.e. [θ] as in tableau below:

(14)

<table>
<thead>
<tr>
<th></th>
<th>*[+distributed]</th>
<th>* [-distributed]</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. s</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>b. θ</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

Under the assumption that *Cor encapsulates the dependent feature hierarchies, the pattern of English obstruent clusters follows directly from the fact that [s] is the least marked coronal fricative. The analysis is shown in the following tableau:

(15)

<table>
<thead>
<tr>
<th>/θpV/</th>
<th>*+dist</th>
<th>*-ant</th>
<th>Ident-Obs-Place</th>
<th>*-dist</th>
<th>*+ant</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. sp</td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>b. fp</td>
<td>*!</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>c. θp</td>
<td>*!</td>
<td></td>
<td>*</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

The tableau shows that with the dis-encapsulated *Cor, candidate (a), which contains a cluster consisting of s+STOP, surfaces as the optimal candidate. Candidates (b) and (c) both fail because of higher * [+distributed].

To conclude, I have shown how the system proposed so far can easily account for the pattern of English obstruent clusters. As I have argued previously,
this is the most common pattern shown across languages. In the analysis I propose this is the preferred pattern cross-linguistically because s+STOP clusters are the best formed of all obstruent clusters with respect to both the continuancy dimension and the place dimension.

4.5 Case Study IV: German

German is also a case of an unmarked system because it only allows coronal fricatives in pre-obstruent position. However, unlike English, German presents the complication of a complementary distribution between [s] and [ʃ] in pre-obstruent position.

4.5.1 The Obstruent System

The German obstruent inventory is given in the chart below. The chart and all the data that follows is based on Hall (1992)

(16)

<table>
<thead>
<tr>
<th></th>
<th>Labial</th>
<th>Alveolar</th>
<th>Alveopalatal</th>
<th>Velar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stops</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simple</td>
<td>p</td>
<td>b</td>
<td>t</td>
<td>d</td>
<td>k</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>g</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>?</td>
</tr>
<tr>
<td><strong>Fricatives</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-sibilant</td>
<td>f</td>
<td>v</td>
<td></td>
<td></td>
<td>ç/x</td>
</tr>
<tr>
<td>Sibilant</td>
<td>s</td>
<td>z</td>
<td>f</td>
<td>ʒ</td>
<td></td>
</tr>
<tr>
<td><strong>Affricates</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-sibilant</td>
<td>pʃ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sibilant</td>
<td>ts</td>
<td>tʃ</td>
<td>dʒ</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
German analysts disagree on whether the sequences \([ts], [t\,\text{ʃ}]\) and \([pf]\) are single phonemic units (James 1969; Wurzel 1980; Hall 1992) or consonant clusters (Moulton 1947; Moulton 1962; Heike 1972; Ungeheuer 1977; Benware 1986). Moulton, for example, argues that there is no reason why only \([ts], [t\,\text{ʃ}]\) and \([pf]\), and not the other stop+fricative sequences of German, i.e. \([ps], [p\,\text{ʃ}]\) and \([ks]\) should be analyzed as affricates. It must be noted, however, that whereas \([ts], [t\,\text{ʃ}]\) and \([pf]\) fully contrast with the other phonemic units of the language, \([ps], [p\,\text{ʃ}]\) and \([ks]\) instead only occur in word-final position or as initial sounds in rare words of foreign origin. Their distribution, therefore, strongly suggests that these two sets of sounds are inherently different. In the analysis that I present \([\text{ts}], [\text{t\,\text{ʃ}}] \) and \([\text{pf}]\) are treated as affricates, whereas \([ps], [p\,\text{ʃ}]\) and \([ks]\) are treated as clusters.

4.5.2 Phonotactics of Onset Obstruent Clusters

In German, onset obstruent clusters are very limited. Of all the obstruent clusters\(^6\)

---

\(^6\) Note that other obstruent clusters occur in German onsets. Sequences of an obstruent followed by \([v]\) as in \([\text{kvark}] \text{quark}\), or \([\text{tsvaɪ}] \text{two}\), are quite common. I assume that these sequences examples of core clusters rather than obstruent clusters, since, as argued for a number of Germanic languages (König and van der Auwera 1994) the segment \([v]\) is best classified as a sonorant.

In addition, a number of obstruent clusters are also found in words of foreign origin, e.g. \([\text{ptolomeːs}] \text{Ptolomy}; [\text{ktenoiːt}] \text{ctenoid}; [\text{kseːɾks}] \text{xerox}; [\text{pʃɔr}] \text{Pshorr (name)}; [\text{spektərəm}] \text{spectrum}; [\text{stiːl}] \text{style}; [\text{sfrəʊ}] \text{sphere}; [\text{stseːɾa}] \text{scene}. As for the reasons given in Chapter 2, these clusters will not be considered representative of the German phonotactics. Their presence in the language can be explained following the model of lexicon stratification described for Modern Greek.
occurring in the language, the clusters that are considered representative of the native phonotactics are only the ones consisting of a coronal fricative and a stop, as in the following chart:

(17)

<table>
<thead>
<tr>
<th></th>
<th>p</th>
<th>t</th>
<th>k</th>
</tr>
</thead>
<tbody>
<tr>
<td>s</td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>f</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
</tbody>
</table>

Representative examples are given in (18) below

(18) a. *skat* skat

   b. *spi:l* game

   c. *t*tant stand

The chart in (17) shows the distribution of the two coronal fricatives in pre-obstruent position. On the basis of the overall distribution of these two segments in the language, many authors (Trubetzkoy 1939; Wurzel 1970; Werner 1972; Scholz 1972; Hall 1992) have pointed out that in German [s] and [f] are nearly in complementary distribution in pre-consonantal position. The generalization, according to these authors, is that [f] occurs before all [-high] consonants\(^7\) in syllable initial position, and [s] occurs elsewhere. The two segments, however, contrast in inter-vocalic and final position. The following chart from Hall (1992),

\(^7\) [-high] consonants in German are: [p b t d v s z m n l r h]
summarizes the relevant facts. In the chart, the parentheses indicate that clusters are rare.

(19)

<table>
<thead>
<tr>
<th></th>
<th>syllable-initial</th>
<th>medial</th>
<th>syllable-final</th>
</tr>
</thead>
<tbody>
<tr>
<td>sp</td>
<td>(+)</td>
<td>+</td>
<td>(+)</td>
</tr>
<tr>
<td>st</td>
<td>(+)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>sk</td>
<td>+</td>
<td>+</td>
<td>(+)</td>
</tr>
<tr>
<td>sv</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sjp</td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>jšt</td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>jšk</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>jšv</td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>sn</td>
<td>(+)</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>sm</td>
<td>(+)</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>sl</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sř</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>jšn</td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>jšm</td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>jšl</td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>jšř</td>
<td></td>
<td>+</td>
<td></td>
</tr>
</tbody>
</table>

On the basis of the patterns above, German can be classified as a *Type 1* language for onset obstruent clusters, i.e. a language that only allows FS clusters. Moreover, on the place dimension, the FS clusters are only limited to sequences of a coronal fricative followed by a stop.
4.5.3 Analysis of Onset Obstruent Clusters

German is an example of a *Type 1* language in onset, i.e. a language that only admits FS clusters. On the continuancy dimension, the language is characterized by the same ranking as English:

\[(20) \quad \text{OCP[-cont]}, \text{OCP[+cont]}, \ast \text{SO} \gg \text{Ident(continuant)}\]

On the place dimension, it has been observed that only the coronal fricatives in the language are allowed in pre-obstruent position. In particular, German shows a complementary distribution between [s] and [ʃ]. A number of researchers (Wurzel 1970; Scholz 1972; Standwell 1973; Hall 1992) have analyzed the distribution of [s] and [ʃ] in syllable-initial position by assuming that [s] is the underlying segment and [ʃ] is derived via some kind of phonological rule.

Following the intuition of these authors and under the assumption that [s] is the unmarked coronal fricative, I also account for the complementary distribution of [s] and [ʃ] in pre-consonantal position, by analyzing words containing [sk] as the normal default case, and words containing [ʃt] and [ʃp] as emerging under the effects of a higher ranked constraint that disqualifies [st] and [sp] as potentially optimal candidates. I propose that this constraint is a
constraint that disallows a sequence consisting of an [s] followed by a [-high] in the onset. The constraint\(^8\) is defined below:

\[(21) \quad (*\mathrm{s}[-\mathrm{high}])^{\mathrm{onset}}:\]

Disallow a sequence consisting of an [s] followed by a [-high] in the onset.

Tableau (22) below contains the relevant sub-hierarchy of *Cor. The tableau shows that an input of the type [ft] would automatically default into a cluster containing [s] because of the fact that [s] is less marked than [f]. However this is a wrong result in German because it does not capture the complementary distribution between the two segments.

(22)

<table>
<thead>
<tr>
<th></th>
<th>Ident-Obs-Place</th>
<th>*[-ant]</th>
<th>*[+ant]</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ft</td>
<td>*</td>
<td>*</td>
<td>**</td>
</tr>
<tr>
<td>b. jt</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

---

\(^8\) Note that although the use of the feature [high] for consonants is obsolete, this seems to be the best way of accounting for the complementarity. Note further that this constraint could not be a conjoined constraint of *s and *[-high] in the onset domain since it would equally fail [st], [sp] and [sk] because the latter would violate the constraint as well. As a matter of fact, although [k] is a [+high] segment, [s] is [-high]. So the [s] in an [sk] cluster would incur violation of the conjoined constraint as well because it is an [s] and it is [-high]. Note also that an OCP on [-high] is not an option because it would otherwise rule out onset clusters such as [pl].
However, defaulting to a cluster whose initial member is [s] is prevented under duress of the constraint in (21), as shown in the tableau below:

(23)

<table>
<thead>
<tr>
<th>/ftV/</th>
<th>*s[-high]</th>
<th>Ident-Obs-Place</th>
<th>*[-ant]</th>
<th>*+[+ant]</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. st</td>
<td>*!</td>
<td>*</td>
<td>*</td>
<td>**</td>
</tr>
<tr>
<td>b. f</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

By dis-encapsulating *Cor, the relative ranking of Ident-Obs-Place and the relevant sub-hierarchies must be reconsidered. The alternation provides evidence that Ident-Obs-Place must be dominated by at least *[-ant] in order to prevent an input of the form [*k] from surfacing faithfully due to the fact that Ident-Obs-Place dominates *[-ant]. This undesired result is illustrated in the following tableau:

(24)

<table>
<thead>
<tr>
<th>/fkV/</th>
<th>*s[-high]</th>
<th>Ident-Obs-Place</th>
<th>*[-ant]</th>
<th>*+[+ant]</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. sk</td>
<td>*!</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>b. f</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

This input shows that the right relative ranking of Ident-Obs-Place and the sub-hierarchy of *Cor must be the one shown in tableau (25) below:
With Ident-Obs-Place dominated by *[-ant], an input of the type in (25) correctly defaults to the least marked fricative. On the other hand, however, higher ranking of *s[-high] will determine that an input of the form [st] cannot make it to the surface, despite the fact that it contains the unmarked fricative. This is shown in tableau (26) below.

(25)

<table>
<thead>
<tr>
<th></th>
<th>*[-high]</th>
<th>*[-ant]</th>
<th>Ident-Obs-Place</th>
<th>*[-ant]</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>sk</td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>b.</td>
<td>†k</td>
<td></td>
<td>*!</td>
<td></td>
</tr>
</tbody>
</table>

Finally, the analysis I propose in this dissertation also accounts for the fact that in medial or in syllable final positions an underlying [st] never surfaces as an [ʃt] cluster. It has been argued (Hall 1992) that in German, as in many Indo-European languages s+Obstruent clusters are syllabified heterosyllabically. In a

---

9 Remember that the contrast in the inventory is maintained by high ranking IdentRelPlace.
word such as “Minister”, then the medial [st] cluster would satisfy the constraint 
*s[-high]Onset because this constraint has been relativized to the onset domain. If 
the cluster is not a tautosyllabic onset cluster, the constraint is vacuously satisfied.

In conclusion, I have shown that in onset position, German is a Type 1 
language, which allows a relatively marked phonotactics for the FS clusters. 
Moreover, I have argued that the allophonic variation between s/ʃ in pre-
consonantal position can be explained by maintaining the unmarkedness of the 
fricative [s]. In other words, initial [sk] is not considered an exception to the more 
regular distribution [ʃt], [ʃp]. In those environments in which [s] is banned to 
surface due to high ranking of the sequential constraint, *s[-high]Onset, the 
occurrence of [ʃ] is the next best choice.

4.6 Case Study V: Onset Place Asymmetries in Delaware

The Delaware language is spoken in Ontario, approximately fifty miles southwest 
of London, Ontario. The Delaware speaking people have migrated over 300 years 
from the Manhattan Island to various locations in the US and Canada. The 
importance of Delaware for the purpose of this typology is that it represents an 
asymmetric system for place in onset obstruent clusters. The language allows all 
types of obstruent clusters in the onset. However, whereas only coronal fricatives 
are allowed in pre-obstruent position, stops can occur at any place of articulation 
in the same position. This is not an isolated example of such a system; Modern
Hebrew is, indeed, another one.

The following chart of the Delaware obstruent system is based on O’Meara J. (1996)\textsuperscript{10}

(27)

<table>
<thead>
<tr>
<th></th>
<th>Labial</th>
<th>Alveol</th>
<th>Alveo-Palatal</th>
<th>Velar</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stops</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simple</td>
<td>p b</td>
<td>t d</td>
<td>k g</td>
<td></td>
</tr>
<tr>
<td><strong>Fricatives</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-sibilant</td>
<td>(f v)</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Sibilant</td>
<td>s z</td>
<td>ʃ z</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Affricates</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sibilant</td>
<td></td>
<td>ċ ž</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The parentheses around the labial fricatives indicate that these sounds are only found in English borrowings, and are therefore not considered part of the native inventory. The author, rather than distinguishing between voiceless/voiced sounds, uses “strong” and “weak” respectively. The weak member of each pair occurs only post-nasally. They never occur word-initially thus suggesting their allophonic status.

As for syllable structure, the language contains both core clusters and obstruent clusters. The onset obstruent cluster phonotactics is quite rich. A chart of two member initial clusters\textsuperscript{11} is given below. All the clusters occur in monomorphemic words.

\textsuperscript{10} The writing system used in this dictionary is intended to enable speakers of Delaware and non-native speakers to read and write Delaware.

\textsuperscript{11} I have excluded [f] from the chart since it is not a native sound and never occurs in clusters.
(28) Two member initial clusters

<table>
<thead>
<tr>
<th>p</th>
<th>t</th>
<th>č</th>
<th>k</th>
<th>s</th>
<th>ʃ</th>
<th>x</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>pt</td>
<td>pk</td>
<td>ps</td>
<td>pʃ</td>
<td>px</td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>tp</td>
<td></td>
<td></td>
<td>tx</td>
<td></td>
<td></td>
</tr>
<tr>
<td>č</td>
<td>čp</td>
<td></td>
<td>čk</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k</td>
<td>kp</td>
<td>kt</td>
<td>kč</td>
<td>ks</td>
<td>kʃ</td>
<td>kx</td>
</tr>
<tr>
<td>s</td>
<td>sp</td>
<td></td>
<td>sk</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ʃ</td>
<td>ʃp</td>
<td>ʃt</td>
<td></td>
<td>ʃx</td>
<td></td>
<td></td>
</tr>
<tr>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The chart shows that clusters whose initial member is a fricative can only contain one of the coronals. There are no clusters whose initial fricative is the velar [x]. On the other hand, there are plenty of clusters whose first member is a stop at places of articulation other than coronal. This asymmetry is explained in the system I propose by assuming the existence of the special faithfulness constraint relative to stops repeated below:

(29) Ident-Stop-Place

Stops must agree in place features with their output correspondent.

The asymmetric behavior of Delaware obstruent clusters then follows from the interaction of this special faithfulness constraint with the rest of the hierarchy. The ability of stops occurring in pre-consonantal position to maintain their input
place of articulation is due to high ranking Ident-Stop-Place as shown in tableau that follows:

\[(30)\]

<table>
<thead>
<tr>
<th>/kpV/</th>
<th>Ident-Stop-Place</th>
<th>*Dor</th>
<th>Ident-Obs-Place</th>
<th>*Cor</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. kpV</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. tpV</td>
<td>*!</td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

Candidate (b), in which the first stop of the cluster neutralizes to coronal place loses in the competition with the faithful candidate (a) because of its violation of Ident-Stop-Place, which requires that stops maintain their input place of articulation regardless of their position.

On the other hand, an input containing a velar fricative as first member of an obstruent cluster will undergo neutralization as shown below:

\[(31)\]

<table>
<thead>
<tr>
<th>/xpV/</th>
<th>Ident-Stop-Place</th>
<th>*Dor</th>
<th>Ident-Obs-Place</th>
<th>*Cor</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. xpV</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. spV</td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

Candidate (31a) loses because *Dor dominates the only faithfulness constraint that applies to a fricative in pre-obstruent position. Due to this domination relation, dorsal place is disallowed in pre-consonantal position unless the segment in that position is a stop, as shown in the previous tableau. Note that the presence
of the Ident-Place-Stop is crucial because in its absence an input containing a dorsal stop in pre-obstruent position would undergo place neutralization as well, as shown in tableau (32) below:

(32)

<table>
<thead>
<tr>
<th>/kpida/</th>
<th>Ident-Rel-Place</th>
<th>*Dor</th>
<th>Ident-Obs-Place</th>
<th>*Cor</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Desired winner:</td>
<td>✤ kpV</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Wrong winner:</td>
<td>☛ tpV</td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

Candidate (a), in which place is maintained in both members of the clusters, loses in the competition with a candidate in which place is neutralized in the first member of the cluster. This is due to the violation of dominant *Dor.

4.7 Case Study VI: Takelma

Takelma is an extinct Penutian American Indian language. The main sources for the data are Sapir (1922) and Borim (1991). Takelma is interesting because it represents a language in which both fricatives and stops in pre-obstruent position are restricted to coronal place. Moreover, due to this place restriction, the language admits only FS and SS clusters, but not SF clusters. So, in other words, the language admits a more marked type of obstruent clusters, i.e. SS, but not a less marked type, i.e. SF. I argue that this surface pattern is the result of a conflict
between the two markedness dimensions relevant to obstruent clusters. This conflict results in a harmonically incomplete system.

4.7.1 The Facts about Takelma

The obstruent system of the language is given in the chart below taken from Borim (1991):

(33)

<table>
<thead>
<tr>
<th></th>
<th>Labial</th>
<th>Coronal</th>
<th>Dorsal</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stops</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lenis</td>
<td>p</td>
<td>t</td>
<td>k</td>
</tr>
<tr>
<td>Fortis</td>
<td>p’</td>
<td>t’</td>
<td>k’</td>
</tr>
<tr>
<td>Aspirated</td>
<td>p^h</td>
<td>t^h</td>
<td>k^h</td>
</tr>
<tr>
<td><strong>Fricatives</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lenis</td>
<td>s</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fortis</td>
<td>ts’</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In Takelma, onset obstruent clusters are very restricted. According to Sapir (1922) common clusters are the ones in (34)^12

(34) FS: sp sk sk^w

*SF

SS: t^hp t^hk t^hk^w


^12 [st] is extremely rare.
As the data in (34) shows, only FS and SS clusters are allowed. There are no SF clusters in the language. However, the initial segment in the obstruent cluster is always a coronal, never a labial or a dorsal. As I will show, this pattern is predicted by the constraint system proposed to capture the place generalizations. There are no clusters with two coronal place segments, or two segments with the same place of articulation. In other words, although its manner features are well-formed, any SF cluster would violate the place restrictions on obstruent clusters.

From the place point of view, the ranking that predicts the pattern of coronal place neutralization for both stops and fricatives in pre-obstruent position is given in (35) blow:

(35)  Ident_{Rel}Place >> *Dor, *Lab >> Ident-Obs-Place, Ident-Stop-Place, *Cor

The tableau that follows shows how an input with a Dorsal is neutralized to coronal place, given the ranking in (35):

(36)

<table>
<thead>
<tr>
<th>/kʰpV/</th>
<th>Ident_{Rel}Place</th>
<th>*Dor</th>
<th>Ident-Obs-Place</th>
<th>Ident-Stop-Place</th>
<th>*Cor</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. kʰpV</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. т³pV</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

Candidate (a) fails because the dorsal consonant is not in a release position and is not protected by high ranking Ident_{Rel}Place. Candidate (b) is the optimal form
because the pre-obstruent stop has the least marked place of articulation, i.e. coronal place.

On the continuant dimension, the fact that SS clusters occur suggests that Takelma has the ranking that predicts that also SF clusters would surface. The relevant ranking is given in (37) below:

(37) OCP [+cont] >> Ident(cont) >> OCP[-cont], *SO

As argued in Chapter 2, this ranking entails that FS, SF and SS surface in the language. However, in the case of Takelma the place hierarchy has priority over the manner hierarchy and thus prevents SF clusters from occurring. The reason why SF cannot occur is that [dorsal] and [labial] are not allowed in pre-obstruent position. This requirement restricts the range of possible SF clusters. Given the pattern in (34), the only possible SF cluster would be \([t^h s]\). However, in Takelma, obstruent clusters are never allowed to share place of articulation. Due to high ranking OCP(Place) an input containing \([t^h s]\) will surface as a cluster with two stops at different places of articulation as shown in the following tableau:

(38)

<table>
<thead>
<tr>
<th>(/t^h s/)</th>
<th>OCP(Place)</th>
<th>Ident(_{Re\text{fPlace}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. (t^h s)</td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>b. st</td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>c. (t^h p)</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>
In conclusion, the interaction of the two markedness hierarchies and the fact that (35) has priority over (37), results in Takelma being *harmonically incomplete*. Although the constraint ranking in (37) allows for both SF and SS to surface, the former are prevented from occurring because of the fact that the place hierarchy has priority over the manner hierarchy.

### 4.7.1.1 An Additional Fact about Takelma

In addition to the fricative [s], Takelma also contains a sound which is transcribed as [x]. According to Sapir (1922), this segment is derived from original [ts’] and behaves phonologically like a coronal. In particular, [xt] and [st] clusters are extremely rare and I assume here that they are ill-formed. I assume that the absence of [xt] clusters is an accidental gap due to the fact that they would only have arisen from the ill-formed [coronal]-[coronal] sequence *[ts’t]*.

In the following section, I will discuss an alternative to the system I have just proposed. The alternative system is based on Steriade’s Licensing-by-Cue. I argue that in the case of place features the system presents a number of typological problems.

### 4.8 Licensing by Cue

Steriade (1997) proposes deriving alternations in sound patterns on the basis of a model of grammar that makes explicit reference to independently known facts about the perception and production of speech. In particular, she proposes that
laryngeal features are neutralized in positions where perceptibility factors are impoverished, and on the contrary, maintained in positions where perceptual cues are strongest. She refers to this hypothesis as “Licensing by Cue”\(^{13}\) and models it on the basis of alignment between constraint hierarchies and harmonic scales. Specifically, in the case of voicing, she assumes a perceptibility scale for voicing based on the contexts in which contrastive voicing is more or less likely to be identified. The scale is given in (39) below. The symbol \( \rightarrow \) indicates that voicing contrast in one phonological context is more perceptible than in the context to its right.

(39) Steriade’s scale of obstruent voicing perceptibility according to context.

\[
V_{(+son)} \rightarrow V_{(#)} \rightarrow V_{(-son)} \rightarrow \{[-son]_{(-son)}, [-son]_{#}, #_{(-son)}\}
\]

The scale expresses the fact that perceptual cues to obstruent voicing are stronger when the obstruent occurs between a vowel and a sonorant and becomes increasingly weaker as we move to the right. In Steriade’s model perceptibility scales of the type in (39) project families of constraints that correspond to each context in the perceptibility scale. Such constraints are formulated as negative constraints banning the occurrence of a feature F in every individual context in the scale. Steriade’s constraints are, in other words, positional markedness constraints. In the case of obstruent voicing, Steriade proposes the family of

---

\(^{13}\) As opposed to models that postulate a correlation between syllabic positions and sites of licensing or neutralization (Licensing by Prosody) (Ito 1986, 1989, Goldsmith 1990, Rubach 1990, Lombardi 1991, 1995).
constraints in (40) that are universally ranked on the basis of the perceptibility scale in (39).

(40) \[ *\alpha_{\text{voice}/[-\text{son}]}[-\text{son}], [-\text{son}]_#, #[-\text{son}] \]

\[ *\alpha_{\text{voice}/V][-\text{son}] \]

\[ *\alpha_{\text{voice}/V}# \]

\[ *\alpha_{\text{voice}/V}[+\text{son}] \]

Patterns of voice neutralizations are derived by the interaction of the fixed hierarchy in (40) with a faithfulness constraint that preserves input voice values, \textit{Preserve [voice]}. If the faithfulness constraint only dominates the lowest ranked constraint \[ *\alpha_{\text{voice}/V}[+\text{son}] \], then voice contrast is only preserved after vowels and before sonorants, i.e. the context in which cues to voicing are strongest. If \textit{Preserve[voice]} dominates both \[ *\alpha_{\text{voice}/V}# \] and \[ *\alpha_{\text{voice}/V}[+\text{son}] \], then voice is licensed in the two corresponding environments, i.e. after vowels and before sonorants or word finally. With \textit{Preserve[voice]} ranked above \[ *\alpha_{\text{voice}/V][-\text{son}] \] voice contrast is maintained also after vowels and before non-sonorant segments. Finally, when \textit{Preserve[voice]} dominates the whole constraint hierarchy, voice contrast is preserved between non-sonorant segments and either preceding or following a non-sonorant, i.e. it is preserved in both strong and weak environments.
A strict implementation of Steriade’s model can be carried out by reference to the so-called contextual or transitional cues. These cues are found in the brief transitional period between a consonant and an adjacent vowel, or a sonorant. Obstruents also have internal cues. Internal cues are found in the burst for stops and in the frication noise for fricatives. They have been shown to be weaker than contextual cues. Therefore, perception of obstruents occurring either before or, to a lesser extent, after a vowel is stronger because the listener has access to both types of cues, i.e. contextual and internal. On the other hand, perception of obstruents in the context of other obstruents is impoverished due to the fact that the listener can only rely on internal cues. For this reason, it is argued that pre-obstruent and post-obstruent positions are the weakest positions where place contrast can be maintained (Steriade 1997, Wright 1996 and references therein). Segments occurring in these positions are therefore more likely to undergo neutralization of place than obstruents occurring in positions adjacent to a vowel.

On the basis of these perceptual facts about obstruents, there are basically four main contexts that need to be identified and on the basis of which a perceptibility scale á la Steriade can be formulated. The contexts in question are:
These four contexts give rise to the following perceptibility scale:

\[(42) \quad _{(S)V} \triangleright V(S)_- \triangleright \{_O, O_-\} \]

The context of a following vowel represents the strongest cues to obstruent place due to the fact that two types of cues are accessible in this context, i.e. contextual and internal. The context of a preceding vowel, and precisely the transitional period between a vowel and a following obstruent, still makes both types of cues available, but they have been shown to be weaker than the ones in pre-vocalic position. Finally the weakest cues are found in contexts where vowel transitions are absent, i.e. in pre-obstruent and post-obstruent positions. Only internal cues are available in such positions. Although four different contexts are represented in the scale, the scale itself basically expresses the main fact that cues to obstruent place are stronger if the obstruent is adjacent to a vowel, because of the presence of vowel transitions, and weaker if the obstruent is adjacent to another obstruent, because vowel transitions are absent in this context.

In Steriade’s model, perceptibility scales project families of markedness constraints that ban a certain feature in each individual context. However, by

---

\[ ^{14} \text{This is the same context given in Lombardi (1991).} \]
projecting markedness constraints for place features, the system of constraints would grow considerably, as shown in (43) below:

(43) Positional Markedness Hierarchies:

(a) *Lab{/O, O/} >> *Lab/V(S)_ >> *Lab/_/S)V
(b) *Dor{/O, O/} >> *Dor/V(S)_ >> *Dor/_/S)V
(c) *Cor{/O, O/} >> *Cor/V(S)_ >> *Cor/_/S)V

Under Prince and Smolensky’s markedness hierarchy, place features are themselves ranked. This is shown below:

(44) *Lab, *Dor >> *Cor

In order to maintain this ranking, the individual hierarchies would need to be ranked accordingly. So in other words, we would have the following ranked hierarchy:

(45) *Lab{/O, O/}, *Dor{/O, O/}
    |   *Lab/V(S)_ , *Dor/V(S)_
    |   *Lab/_/S)V , *Dor/_/S)V
    |   *Cor{/O, O/}
    |   *Cor/V(S)_
    |   *Cor/_/S)V
This positional markedness hierarchy, with the interaction of a general faithfulness constraint that preserves place, e.g. \textit{Preserve[Place]} has a number of problems. If \textit{Preserve[place]} is ranked anywhere above the \textit{*Cor} family, the system predicts that Coronal place is allowed in all three positions that have been identified, i.e. presonorant, coda and pre-obstruent position. Similarly, the system predicts that if [labial] is allowed in one context [dorsal] is also, and vice versa. So for example, with \textit{Preserve[Place]} dominating \textit{*Lab/[\_O, \_O]}, \textit{*Dor/[\_O, \_O]} we are not only predicting that both [labial] and [dorsal] place are allowed in pre-obstruent position but also that both must necessarily occur in coda. Therefore, this system does not allow us to freely rerank \textit{*Dor/V(S)} because the context is fixed in the hierarchy of contexts. This predicts a correlation between the patterns in clusters and in codas, but there is no evidence for such a correlation.

Similar problems arise if we assume that, rather than a single \textit{Preserve[Place]}, the system consists of three different faithfulness constraints relativized to the different place features. The ranking among the three different faithfulness constraints would also be fixed on the basis of the hierarchy in (44):

\begin{equation}
\text{(46)} \quad \textit{Preserve[coronal]} >> \textit{Preserve[labial], Preserve[dorsal]}
\end{equation}

Interaction of (46) with (45) would also predict that if [labial] and [dorsal] place are allowed in pre-obstruent position, they would also necessarily occur in coda because of the fixed ranking among the faithfulness constraints.
Another possibility would be that (46) actually interacts with a hierarchy of the type in (47) below:

(47) \( \textit{Place/\{_O, O_\}} \gg \textit{Place/V(S)}_\gg \textit{Place/\_}(S)V \)

Also in this case, the system would predict that if [dorsal] or [labial] is allowed in pre-obstruent position, i.e. \( \textit{Preserve[labial]} \gg \textit{Place/\{_O, O_\}}, \) it must necessarily be allowed in coda position as well, because \( \textit{Preserve[labial]} \gg \textit{Place/\{_O, O_\}} \) implies \( \textit{Preserve[labial]} \gg \textit{Place/V(S)}_\). Therefore, no matter how we implement Steriade’s model, a correlation between place in pre-obstruent position and coda position is always predicted.
5.1 Introduction

In the previous chapters I have argued that different degrees of markedness exist among the various obstruent clusters. I have shown that, on the various dimensions relevant to obstruent clusters, s+STOP onsets turn out to be the least marked of all the clusters of this type. In particular, they are unmarked along the dimension of the feature [continuant] because they are FS clusters. On the place dimension, they are the least marked of all because the fricative in pre-obstruent position is a coronal.

The main purpose of this final chapter is to derive the unmarkedness of s+STOP clusters from the constraint system proposed and to defend the view that s+STOP onsets are indeed regular tautosyllabic onset clusters rather than heterosyllabic or monosegmental, as argued by a number of authors. I will show how the strongest pieces of evidence that have been used in support of the view that s+STOP onsets are not regular tautosyllabic onsets can find independent explanations. In particular, I argue that the heterosyllabicity of medial s+STOP in Italian does not provide evidence for the ill-formedness of these clusters in the language, as argued in previous approaches. I show that such a pattern of syllabification is just a consequence of minimal violations of basic OT syllable structure constraints. I also discuss the case of Sanskit reduplication and argue
that this process as well cannot be considered evidence for the ill-formedness of s+STOP clusters in this language.

5.2 s+STOP Onsets and the Sonority Sequencing Principle

Onset clusters consisting of s+STOP constitute a major outstanding problem in previous phonological theories. Such clusters constitute violations of the Sonority Sequencing Principle (see discussion in Chapter 1). Depending on whether fricatives and stops are assigned the same or different sonority values on the sonority scale, these clusters represent either "sonority plateaus" or "sonority reversals", respectively. But despite the fact that they violate the SSP under any version of the sonority scale, these types of clusters are quite common among a significant number of languages that allow complex syllable margins. As a matter of fact, a common phonotactic pattern in the onset consists of core clusters, i.e. obstruent+sonorant, and s+STOP clusters. In other words, in many systems the only obstruent clusters allowed are s+STOP clusters.

In order to reconcile their cross-linguistic occurrence and the SSP, many researchers have proposed that s+STOP clusters enjoy a special status in phonological theory and are therefore immune to the principle. Among the many proposals, for example, Steriade (1982, 1988) and Clements (1990) argue that these clusters are created by post-cyclic syllabification rules, which are not constrained by relative sonority. They argue that their special status lies in the fact that, at the level where the SSP is relevant, i.e. core syllabification, these clusters
do not form tautosyllabic sequences. Their heterosyllabic at that point makes them immune to the SSP because the principle holds over tautosyllabic sequences only. For Harris (1994), clusters of this type are never tautosyllabic. He argues that the /s/ in an s+STOP cluster is not part of the onset, but it rather belongs to the coda of a preceding syllable or of a nucleusless syllable, in the case of initial onsets. Other researchers, such as Fudge (1969) and Selkirk (1982), propose that their special behavior lies in the fact that they are single onsets, thus able to escape the SSP because the SSP holds over tautosyllabic clusters and not mono-segments. From a representational point of view, Broselow (1991) proposes that the fricative in an s+STOP onset is licensed by virtue of its link to the following stop (which she calls "parasitic licensing"). In the same vein, Steriade (1994) proposes that their mono-segmental status is structurally represented by the fact that, in an s+STOP sequence, the fricative does not occupy an independent position but rather a segment internal slot. She defines such a position as an Approach-to-Closure position. Finally, Fujimura (1995, 1996, 1997), within his Converter/Distributor Model of phonetic implementation, analyzes s+STOP clusters as integral units specified with the feature {spirantized}. The feature {spirantized} is implemented by two concurrent elemental gestures, one for the frication generation and the other a stop closure.

The main problem with such approaches is that they stipulate special syllabification rules or representations for s+STOP clusters in order to justify their immunity to the SSP. However, these clusters are quite common clusters, which
makes an approach that treats them as “exceptions” not satisfactory. The typology that I propose shows that these types of clusters are indeed the unmarked obstruent clusters. Treating them as marked, i.e. exceptions, rather than as unmarked phonotactics reflects a misunderstanding of the facts themselves. Moreover, strong independent evidence to support their "special" status with respect to the SSP is not always easy to find. And when such evidence can be found, it can, indeed, be easily attributed to independent principles of grammar, as I will show later in the chapter. Finally, I will argue that even one of the strongest pieces of evidence found in Italian for the heterosyllabicit of initial s+STOP clusters follows straightforwardly from independent facts. Specifically, I show how the syllabification of medial s+STOP clusters and the allomorphic alternation of the definite article in Italian is just a natural consequence of the ranking of basic markedness constraints in Optimality Theory (Prince and Smolensky 1993).

5.3 s+STOP Onsets: the Unmarked Phonotactics

Based on the results obtained from the typological study introduced in the previous chapters, we see that s+STOP clusters are the least marked obstruent clusters, and therefore not exceptional as suggested in previous literature. I argue that these types of clusters form regular onset tautosyllabic clusters, just like clusters such as /tr/ do. Within their own domain, onset s+STOP clusters are the unmarked cluster type by virtue of being unmarked along both the manner and the place dimensions.
On the manner dimension, s+STOP clusters are unmarked because they consist of a fricative followed by a stop. I have argued, in Chapter 2, that sequences of this type are unmarked along the continuancy dimension. Recall that their unmarkedness results from the fact that FS clusters do not violate any of the constraints proposed in Chapter 2. I repeat the tableau showing their unmarkedness below:

(1)

<table>
<thead>
<tr>
<th></th>
<th>OCP[+cont]</th>
<th>*SO</th>
<th>OCP[-cont]</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>FS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>FF</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td>SF</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>d.</td>
<td>SS</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

The tableau shows that, under this set of constraints, FS is the most harmonic of all the clusters because it receives no marks on any of the constraints. For this reason, it is claimed to be the unmarked cluster among the obstruent clusters.

On the place dimension, s+STOP clusters are the most harmonic clusters because the least marked place surfaces in pre-obstruent position. I have argued, in Chapter 4, that in this position place distinctions are harder to maintain due to the fact that this is not a release position and perceptual cues are impoverished. Due to absence of strong place cues, the fricative in pre-obstruent position is most likely to undergo place neutralization. From this view, it follows that the least marked onset obstruent systems are systems in which only coronal, i.e. the least
marked place of articulation, is licensed in a position of weak perceptibility. The ranking that formalizes this intuition corresponds to a ranking of place neutralization in a position where transitional cues are absent. The ranking is repeated below:

(2) Ranking for place neutralization in pre-obstruent position

\[ \text{Ident}_{\text{RelPlace}} \gg *\text{Lab} \gg \text{Ident-Obs-Place} \gg *\text{Cor} \]

The effect of the ranking in (2) is shown in tableau in (3). A hypothetical input of the form /fp/ which contains a labial in both pre-obstruent and pre-vocalic positions surfaces as /sp/ with a coronal in pre-consonantal position instead.

(3)

<table>
<thead>
<tr>
<th>/fpV/</th>
<th>Ident_{\text{RelPlace}}</th>
<th>*Lab</th>
<th>Ident-Obs-Place</th>
<th>*Cor</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>☞ spV</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>b.</td>
<td>☞ fpV</td>
<td>**!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td>☞ ftV</td>
<td>*!</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>d.</td>
<td>☞ st</td>
<td>*!</td>
<td>**</td>
<td>**</td>
</tr>
</tbody>
</table>

Candidates (c) and (d) both incur a violation of the higher ranked positional faithfulness constraint because the prevocalic /p/ is neutralized to /t/, thus violating the constraint that preserves place contrast in stronger positions. Candidate (b), instead, satisfies Ident-Obs-Place, but violates *Lab twice, once because of the labial fricative and the other because of the labial stop. Notice, however, that the fatal violation is actually the one incurred because of the labial fricative and not the one incurred because of the labial stop, which is incurred by
the winning candidate as well. Parsing a labial in prevocalic position is better than neutralizing it to coronal in order to avoid violation of the higher ranked Ident_{Rel}Place constraint. On the contrary, parsing a labial in pre-obstruent position is worst than neutralizing it because place in that position is only preserved by the general constraint Ident-Obs-Place, which is however dominated by the *Lab constraint. Candidate (a), in which place is preserved in the segment occurring adjacent to a vowel, but neutralized in pre-consonantal position, is the most harmonic candidate because it only incurs violations of lower ranked constraints.

Under the proposal that treats s+STOP clusters as unmarked clusters within their own domain, the fact that they are quite common cross-linguistically is the logical consequence of the analysis and does not need to be stipulated. A pattern in which core clusters and s+STOP clusters occur in the onset is, therefore, a relatively unmarked system. Although s+STOP clusters are relatively more marked than core clusters, they are, however, the least marked among the obstruent clusters.

In this section I have discussed the fact that s+STOP onsets clusters represent the least marked among the obstruent clusters. In the next section, I show how some of the evidence that has led previous researchers to argue for the special status of s+STOP clusters can actually be attributed to other principles of grammar. Firstly, I will discuss syllabification of medial s+STOP sequences in Italian. I will show that, contrary to previous analyses, heterosyllabicity of medial clusters does not constitute evidence for the ill-formedness of s+STOP clusters in
the language. On the contrary, it is the result of the interaction of the basic markedness constraints on syllable structure. Secondly, I show how the Sanskrit reduplication patterns discussed in Steriade (1988) are actually a pure case of The Emergence of The Unmarked (McCarthy and Prince 1995) rather than evidence for the ill-formedness of s+STOP clusters.

5.4 Case Study VI: Standard Italian

Standard Italian has been argued to provide one of the strongest pieces of evidence for the special status of s+STOP clusters. In particular, Italian provides clear evidence that in medial position s+STOP clusters are never tautosyllabic. Moreover, s+STOP and core clusters behave differently in certain morphological contexts. These facts have led many researchers (Chierchia (1983); Kaye, Lowenstamm and Vergnaud (1990); Burzio (1989); Davis (1990); Harris (1994)) to argue that s+STOP clusters are not formed by the regular phonotactic rules, but rather by later adjunction rules. In other words, for these authors s+STOP clusters are ill-formed at the level of core syllabification and become well-formed at a later stage of the derivation. I propose, instead, that these facts have an independent explanation and cannot be considered evidence for the ill-formedness of s+STOP clusters.
5.4.1 Cluster Phonotactics

In this section I will provide a description of Italian cluster phonotactics. A correct characterization of what can be considered representative of the native phonotactics is necessary in order to properly characterize the grammar as well.

Consonant clusters in Italian can be grouped into the following three major categories:

(4) a. Core clusters (pr, pl, br, bl, tr, dr, kr, gr, gl, fr, fl)\(^1\)

b. Obstruent clusters (FS: sp, st, sk)\(^2\)

c. s+sonorant/fricative clusters (sm, sn, sl, sr, sf)

Core clusters include all the clusters that contain any obstruent (except /s/) followed by a liquid. I consider these, together with the obstruent cluster of the form s+STOP, the core phonotactics of the language. Both types show the same distributional properties. In initial position, they are the only ones that are also found in monomorphemic words. Both types occur in initial and medial position in monomorphemic words. I provide examples below:

\(^1\) The cluster [pn] is only found in a few words of Greek origin and is therefore excluded as an example of native phonotactics.

\(^2\) Note that in Italian s/z are allophones of the same phoneme and that s+STOP clusters must agree in voicing. Clusters such as [zb], [zd] and [zg] also occur. [ps] and [ks] are also found in a few words of Greek origin and again are not considered part of the native phonotactics
The clusters of the type in (4c), on the other hand, all consist of /s/ followed by either a sonorant or a fricative. The vast majority of s+C clusters is indeed formed by affixation of the prefix /s-/.

This is an extremely productive prefix that can attach to verbs, nouns and adjectives and has different functions depending on the lexical category of the stem. Clusters such as [sm, sn, sl, sr], although well-formed with respect to the SSP, cannot be considered representative examples of core clusters because they do not show the same
distributional properties of core clusters. Unlike the latter, these clusters occur in initial and medial position only in borrowed words or morphologically complex words. A few representative examples are given below:

(6) smeraldo (Latinate\textsuperscript{4})  “emerald”
    smalto (French smalt)  “enamel”
    asma (Latinate)  “asthma”
    sleale (s+leale)  “not loyal”
    islamico (foreign)  “Islamic”
    sregolato (s+regola)  “immoderate”

Similarly, s+fricative clusters, I believe, cannot be considered core obstruent clusters either because they also only occur in morphologically complex words and are never found in medial position in simplex words with a very few exceptions of words of latinate origin. A few representative examples are given below:

(7) sfera (latinate)  “sphere”
    sfarzo (Neapolitan)  “pomp”
    sfoglia (s+foglia)  “layer”
    asfalto (latinate)  “asphalt”
    asfissia (latinate)  “asphyxia”

The following chart recapitulates the distributional properties of the four types of clusters discussed in monomorphemic words of the native vocabulary.

\textsuperscript{4} I use Latinate for words of both Latin or Classical Greek origin.
Chart (8) clearly shows that s+STOP clusters share the same distributional properties of Core Clusters and can therefore safely be considered regular phonotactics. On the other hand, s+Fricative and s+Sonorant do not share any of the distributional properties of Core Clusters. As shown above, these clusters only occur in borrowings or morphologically complex words. For this reason, I believe they cannot be considered to satisfy the regular phonotactics of Italian.

In conclusion, Italian allows only one type of obstruent clusters, i.e. the unmarked FS s+STOP. FF clusters, although present in the language, are not considered well-formed in terms of the system of constraint on the manner dimension. The clusters that do occur but are not considered regular phonotactics occur because of some other constraints preserving morphological information or non-native phonotactics.

5.4.2 Syllabification

In the previous section, I have argued that only two types of clusters can be considered representative of the native phonotactics, i.e. core clusters consisting of an obstruent (except /s/) and a sonorant, and s+STOP clusters. It has been
shown that these two groups of clusters behave differently phonologically. In medial position, s+STOP clusters\(^5\), but not core clusters, are unambiguously syllabified heterosyllabically, with the /s/ in the coda of the preceding syllable and the second member of the sequence in the onset of the following syllable (Chierchia (1983); Kaye, Lowenstamm and Vergnaud (1990); Burzio (1989); Davis (1990); Harris (1994)). There are two main arguments that favor the heterosyllabicity of medial s+STOP clusters. The first argument comes from vowel length and the second from phonotactics patterns.

In Italian, vowel length is predictable. Stressed vowels in open syllables are lengthened:

\begin{align*}
(9) & \text{a. fa:.to} \quad \text{“fate”} & \text{b. [k]a:pra} \quad \text{“goat”} \\
& \text{me:.ro} \quad \text{“pure”} & \text{sa:.[k]ro} \quad \text{“sacred”} \\
& \text{pe:.lo} \quad \text{“hair”} & \text{re:.tro} \quad \text{“behind”}
\end{align*}

On the other hand, stressed vowels in closed syllables are short:

\begin{align*}
(10) & \text{fat.to} \quad \text{“fact”} \\
& \text{man.to} \quad \text{“coat”}
\end{align*}

Since stressed vowels preceding an s+STOP cluster are systematically short, these clusters must not form complex onsets in medial position, but rather /s/ must close the preceding syllable, as in (11).

\begin{align*}
(11) & \text{pas.ta} \quad \text{“pasta”} \\
& \text{ves.pa} \quad \text{“wasp”} \\
& \text{mos.[k]a} \quad \text{“fly”}
\end{align*}

\(^5\) This process affects also s+Consonant clusters.
The second argument is based on phonotactics considerations, and, in my opinion, is not as strong as vowel length. Chierchia argues that there are no monomorphemic words in Italian such as *pelsto or *persto. He attributes this gap to the fact that, in Italian, there is only one post-nucleic position in the rime. The gap confirms the ill-formedness of s+STOP clusters because syllabification in the case of this words would be impossible. If, indeed, tautosyllabic s+STOP clusters were well-formed, there would be no problem syllabifying *persto as *per.sto. It must be pointed out, however, that a few words containing a medial sequences of a sonorant followed by an s+STOP do indeed occur. Some representative examples are given in (12)

(a) perspicace “acute”
(b) pe(r)spirare “exhale”
(c) co(n)stare “consist of”
(d) co(n)statare “notice”

Except for (12a), all other examples are derived from words of Latinate origin, that were morphologically complex in the source language. Parenthesis around the initial sonorant of the sequence indicates that both forms exist in the language6. A word such as perspicace, therefore can only be syllabified as per.spicace given the fact that Italian codas only allow one post-nucleic position. Based on this data, I believe, this second argument is not as strong as the vowel length argument. I argue, therefore, that the presence of words such as the ones in

6 In my speech, I prefer the full form.
(12) supports the view that s+STOP clusters are not ill-formed in the language, regardless of their heterosyllabic property in (11). I will show that their medial syllabification follows from independent principles.

Based on the fact that in medial position s+STOP clusters are never tautosyllabic, Chierchia (1983), Kaye, Lowenstamm and Vergnaud (1990), Burzio (1989), Davis (1990) and Harris (1994) have argued against the well-formedness of initial s+STOP onsets as well. Chierchia (1983), for example, assumes that /s/ is a stray consonant word initially throughout the word phonology. A later adjunction rule incorporates it into the onset and creates an s+STOP cluster.

I argue that medial syllabification is not evidence for the ill-formedness of s+STOP onsets, but is rather an effect derived from the interaction of basic syllable structure constraints. Unlike previous analyses, in which vowel length was considered evidence for the ill-formedness of initial s+STOP clusters, in the analysis I present, vowel length is not taken to constitute evidence for the fact that these clusters do not form well-formed onsets. The basic insight of the analysis I will present here is that whereas in medial position it is better to syllabify s+STOP as for example [s.t] rather than [.st] in terms of syllable structure, in initial position s+STOP clusters can only be syllabified as [.st] because, in Italian, deletion and epenthesis are not possible options. This intuition is formalized in the following tableau:
Candidate (b), a candidate in which /s/ is syllabified in the coda of a nucleusless syllable, fails because it violates a constraint that requires that all syllables have a nucleus (NUCLEUS). Candidates (c) and (d) fail because repairing a cluster that violates the SSP is worse than allowing it to surface given the ranking MAX-IO, DEP-IO >> SSP.

Under the assumption that s+STOP clusters constitute well-formed onsets, I will explain the fact that they are unambiguously heterosyllabic in medial position in terms of best satisfaction of basic syllable structure constraints. In previous accounts the medial syllabification of these clusters followed from the fact that s+STOP onsets were actually disallowed in the language. Their occurrence word initially was explained in terms of special representations or post-lexical syllabification rules. In the present proposal, I show that there is no need to stipulate that s+STOP clusters are formed by special rules. I argue that they are regular well-formed onsets. Their medial syllabification just follows from

---

For simplicity, I am using the portmanteau constraint SSP in place of *Plateau and *Reversal.
minimal violations of independently motivated constraints on syllable structure.

The constraints that determine the syllabification of medial s+obstruent clusters are given in (14) below:

(14)

- **COMPLEX**
  
  No more than one C may associate to the onset or coda node

- **NOCODA**
  
  Codas are disallowed

Given any input containing a medial sequence as in “pasta”, syllabification results from the interaction of *COMPLEX and NOCODA as shown in the following tableau:

(15)

<table>
<thead>
<tr>
<th>/pasta/</th>
<th>*COMPLEX</th>
<th>NOCODA</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. pas.ta</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>b. pa.sta</td>
<td>*!</td>
<td></td>
</tr>
</tbody>
</table>

Candidate (a), in which the /s/ is syllabified in the coda of the preceding syllable, wins over candidate (b), in which the /st/ cluster is syllabified tautosyllabically, because candidate (a) only violates NOCODA, but satisfies higher ranked *COMPLEX. These constraints must be low ranked in the constraint hierarchy

---

8 Clements (1997) argues that this constraint is actually two separate constraints *COMPLEXCODA and *COMPLEXOnset. Also recall that for me *COMPLEX is indeed a portmanteau constraint.
of Italian, given the fact that both complex syllable positions as well as codas are allowed to surface in the first place.

In order to ensure that clusters that obey the SSP are syllabified as tautosyllabic onset clusters, the two syllable structure constraint must be crucially dominated by a third constraint. The constraint at stake is the Syllable-Contact-Law Constraint (Murray and Venneman 1983; Clements 1990; Hironymous 1999). The constraint must crucially be formulated as a negative constraint in the following way:

(16) Syllable-Contact-Law

A coda must not be lower in sonority than the following onset.

Interaction of Syllable-Contact-Law with *COMPLEX accounts for the syllabification of medial clusters that obey the SSP both in the case of an initial stop and fricative, as shown in the tableaux below:

(17)

<table>
<thead>
<tr>
<th>/metro/</th>
<th>Syll-Contact-Law</th>
<th>*COMPLEX</th>
<th>NOCODA</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ɛɛɛɛ me:.tro</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>b. met.ro</td>
<td>*!</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>
As stated, the constraint applies non-vacuously only if the segment in the coda is lower in sonority than the segment of the following onset, as shown in tableaux (17) and (18). On the contrary, the constraint is vacuously satisfied if the segment in the coda is higher in sonority than the one in the following onset, as shown in the following tableau:

(19)

<table>
<thead>
<tr>
<th>/sarto/</th>
<th>Syll-Contact-Law</th>
<th>*COMPLEX</th>
<th>NOCODA</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. sar.to</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>b. sa.rto</td>
<td></td>
<td>*!</td>
<td></td>
</tr>
</tbody>
</table>

The constraint is also vacuously satisfied if the segments are equal in sonority as in the case of s+STOP clusters. Syllabification of these types of clusters than will depend exclusively on lower ranked *COMPLEX and NOCODA, as shown in (20) below:
Finally, this analysis so far incorrectly forces syllabification of medial s+Sonorant sequences into tautosyllabic onset clusters, because /s/ is lower in sonority than any sonorant segment in Italian.

This is not a correct syllabification for a word such as “asma” because of the fact that the initial vowel is short, which again suggests that these clusters also are syllabified heterosyllabically. In section 5.4.1, I have argued that s+Sonorant clusters cannot be considered well-formed clusters according to the regular phonotactics of Italian due to the fact that they are only allowed in borrowing or morphologically complex clusters. Their ill-formedness must therefore correspond to some constraint, or constraint system presumably having to do with sonority distance, that bans their occurrence. I will informally call this constraint *s+Sonorant. However, I believe that an explanation for the pattern found in
Italian lies somewhere in the right evaluation metric for relative well-formedness among core clusters, which is outside the scope of this dissertation. The existence of this constraint then will force heterosyllabicity in these clusters:

\[(22)\]

<table>
<thead>
<tr>
<th>/asma/</th>
<th>*s+Sonorant</th>
<th>Syll-Contact-Law</th>
<th>*COMPLEX</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.  \textsuperscript{\textdegree} as.ma</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. a.sma</td>
<td>*!</td>
<td></td>
<td>*!</td>
</tr>
</tbody>
</table>

In conclusion I have shown that medial syllabification is not evidence for the ill-formedness of \(s+\text{STOP}\) clusters. On the contrary, their medial syllabification is only the result of the interaction of basic syllable structure constraints.

**5.4.3 Morphological Alternations**

Another piece of evidence that has traditionally been used in support of the claim that initial \(s+\text{STOP}\) clusters do not form onset clusters, is the choice of the masculine definite article allomorph\(^9\). This morpheme in Italian has two forms *il* and *lo*. The allomorph *il* can be considered to be the default case since it occurs with words starting in single consonants or clusters that obey the SSP. The allomorph *lo*, instead, occurs with words starting with either \(s+\text{STOP}\) clusters\(^{10}\),

\(^9\) See Burzio (1989) for a more comprehensive list of elements that show similar alternations between \(s+\text{STOP}\) clusters and core clusters.

\(^{10}\) Indeed \(s+\text{consonant}\).
or one of /k, n, ñ, ts, dz/, which have been argued to be underlyingly long
segments (Chierchia 1983) because they do not show the same alternations as the
other consonants, or a vowel\(^\text{11}\). The alternation is shown in (23) below:

(23)

<table>
<thead>
<tr>
<th>Distribution of il</th>
<th>Distribution of lo</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Single consonants:</strong></td>
<td><strong>Long consonants:</strong></td>
</tr>
<tr>
<td>il ponte the bridge</td>
<td>lo [ʃ]iopero the strike</td>
</tr>
<tr>
<td>il topo the mouse</td>
<td>lo [tʃ]io the uncle</td>
</tr>
<tr>
<td>il [k]orpo the body</td>
<td>lo [ts]aino the knapsack</td>
</tr>
<tr>
<td>il bagno the bath</td>
<td>lo [ŋ]occo the dumpling</td>
</tr>
<tr>
<td>il dente the tooth</td>
<td></td>
</tr>
<tr>
<td>il gatto the cat</td>
<td></td>
</tr>
<tr>
<td>il [tʃ]ielo the sky</td>
<td></td>
</tr>
<tr>
<td>il [dʒ]iorno the day</td>
<td></td>
</tr>
<tr>
<td>il forno the oven</td>
<td></td>
</tr>
<tr>
<td>il volo the flight</td>
<td></td>
</tr>
<tr>
<td>il segno the sign</td>
<td></td>
</tr>
<tr>
<td>il mondo the world</td>
<td></td>
</tr>
<tr>
<td>il nome the name</td>
<td></td>
</tr>
<tr>
<td>il ladro the thief</td>
<td></td>
</tr>
<tr>
<td>il regalo the gift</td>
<td></td>
</tr>
<tr>
<td><strong>Core clusters</strong></td>
<td><strong>Other clusters</strong></td>
</tr>
<tr>
<td>il proposito the purpose</td>
<td>lo stato the state</td>
</tr>
<tr>
<td>il plotone the platoon</td>
<td>lo sforzo the stress</td>
</tr>
<tr>
<td>il treno the train</td>
<td>lo sposo the groom</td>
</tr>
<tr>
<td>il drappo the cloth</td>
<td>lo smeraldo the emerald</td>
</tr>
<tr>
<td>il [k]ranio the skull</td>
<td></td>
</tr>
</tbody>
</table>

\(^{11}\) I will only discuss the alternation in the case of words beginning with clusters, since the form *lo* is truncated before vowels, e.g. lo+ozio \(\rightarrow\) lozio.
From the distribution of the two allomorphs, it can be argued that *il is the default allomorph, because it occurs with single consonant onsets and core clusters, whereas *lo is the special case because it occurs in a more restricted set of environments. In particular it occurs with the long consonants and s+C clusters. I will argue therefore that there is a markedness relationship between the two allomorphs based on their distributional properties which makes *il the least marked of the two by virtue of being the default allomorph.

In Chierchia’s analysis, the selection of *lo rather than *il before words such as the ones in the right-hand column follows from the assumption that the /s/ in the initial s+STOP clusters remains stray until the coda rule syllabifies it with the preceding rime, after incorporation of the article. Since Italian allows only one postnucleic position, it is clear why [los.tato] rather than *[ils.tato] is the correct form.

In the analysis I propose, on the contrary, the choice of the allomorph *lo in words that begin with s+STOP clusters follows straightforwardly from the same constraint interaction that forces a medial s+STOP cluster to be syllabified heterosyllabically. This is illustrated in the following tableau:
Candidates (a) through (c) are all out because they violate dominant *COMPLEX, in particular candidate (a) contains a complex coda whereas both (b) and (c) contain a complex onset. Candidate (d) is optimal because it minimally violates NOCODA.

However, an analysis based solely on the above syllable structure constraints, Syll-Contact-Law, *COMPLEX and NOCODA, would favor the selection of lo also in the context where il should instead appear. This is shown in tableau (25) below.

(25)
As the tableau shows, the Syll-Contact-Law immediately eliminates the two candidates in which a cluster /tr/ is syllabified heterosyllabically. Candidate (c), which is the desired winner, however, loses in the competition with candidate (d) because of its violation of NOCODA. An analysis based solely on syllable structure constraints penalizes the default allomorph il due to the fact that, in terms of syllable structure, il is more marked than lo. lo has the unmarked syllable structure CV, whether il both lacks an onset and contains a coda consonant and has, therefore, the more marked syllable structure VC. In a purely phonological analysis, there is no apparent explanation for why il, rather than lo, is the default allomorph and, in particular, for why il, rather than lo, should occur with words beginning with single consonants, as the data in (26) shows:

(26)  il.ti.po *lo.ti.po  the type
     il.na.so *lo.na.so  the nose
     il.sa.le *lo.sa.le  the salt
     il.\text{[dʒ]}e.lo *lo.\text{[dʒ]}e.lo  the ice

As a matter of fact, the occurrence of il with words such as the ones in (26) creates a more marked syllable structure than what lo would create with the same words. This is shown in tableau (27), where I evaluate the two candidates against the syllable structure constraints, ONSET, which requires that all syllables have an onset, and NOCODA.
Tableau (27) shows that candidate (b), the desired winner, incurs violations of both ONSET and NOCODA, and is, therefore, a more marked candidate than (b) which instead satisfies both syllable structure constraints.

The data in (26) can only be explained by assuming a markedness relationship between the two allomorphs based on their distribution\(^{12}\). From this point of view, *il* is the unmarked form of the definite article, by virtue of being the default allomorph, and *lo* is the marked one since it is the special case. This markedness relationship is implemented via the relative ranking \(*lo \gg *il\), in which the markedness constraint corresponding to the default allomorph is lower ranked than the constraint corresponding to the predictable one. By ranking this sub-hierarchy between *COMPLEX and NOCODA*, the right surface form in the case of a stem beginning with a core cluster is predicted. In tableau (28) below, I

\(^{12}\) J. McCarthy developed this idea in the 1993 seminar at University of Massachusetts. The idea was reported to me by L. Benua (p.c.). Davidson (1999) also explores this idea.
only show the relevant part of the hierarchy:

(28)

<table>
<thead>
<tr>
<th>/DEF+treno/</th>
<th>Syll-Contact-Law</th>
<th>*COMPLEX</th>
<th>*lo</th>
<th>NOCODA</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ilt.re.no</td>
<td>*!</td>
<td>*</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>b. lot.re.no</td>
<td>*!</td>
<td>*</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>c. ⇢ il.tre.no</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>d. lo.tre.no</td>
<td>*</td>
<td>*</td>
<td></td>
<td>*!</td>
</tr>
</tbody>
</table>

Candidates (c) and (d) both incur a violation of *COMPLEX, however, candidate (c) contains the unmarked form of the definite morpheme and wins regardless of its NOCODA violation.

This also predicts the right alternation in the case of stems beginning in single consonants and with s+STOP clusters, as shown in tableaux (29) and (30) respectively.

(29) DEF + words beginning in simple onsets

<table>
<thead>
<tr>
<th>/DEF+tipo/</th>
<th>*lo</th>
<th>*il</th>
<th>ONSET</th>
<th>NOCODA</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. lo.ti.po</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. ⇢ il.ti.po</td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

Tableau (29) shows that when *COMPLEX is not at issue, the relative markedness of the allomorphs determines the choice of the least marked of the two, i.e. *il. However, when higher ranked *COMPLEX is at stake, it forces the
choice of the more marked allomorph lo, which gives a less marked syllable structure. This is shown in tableau (30) below.

(30) DEF + words beginning with an s+Stop cluster

<table>
<thead>
<tr>
<th>/DEF+stato/</th>
<th>*COMPLEX</th>
<th>*lo</th>
<th>*il</th>
<th>NOCODA</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.  ils.ta.to</td>
<td>!</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>b.  il.sta.to</td>
<td>!</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>c.  lo.sta.to</td>
<td>!</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d.  los.ta.to</td>
<td></td>
<td>*</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

In this section I have shown that the different syllabification patterns of s+STOP clusters in initial and medial positions does not necessarily imply that s+STOP clusters are ill-formed onsets and require, therefore, special rules or representations to account for their occurrence in initial position. I show, on the contrary, that the different syllabification patterns follow straightforwardly from the constraint ranking of the language.

So far, I have argued that s+STOP onsets are the best-formed clusters among all the obstruent clusters because they are doubly unmarked. They are unmarked along the dimension where the feature [continuant] is relevant and unmarked along the place dimension, because they surface with the least marked place in a position where place contrast is harder to maintain. Moreover, I have shown that one of the strongest pieces of evidence in support of the ill-formedness of initial s+STOP clusters, i.e. their syllabification in medial position
and the masculine definite article alternation in Italian, can be explained in terms of minimal violations of basic syllable structure constraints. i.e. *COMPLEX and NOCODA, and, does therefore not constitute evidence for the ill-formedness of s+STOP onsets.

In the next section I will examine the case of Sanskrit reduplication and show that this process also is not evidence for the fact that s+STOP clusters are ill-formed in the language.
5.5 Sanskrit Perfect Reduplication

Another piece of evidence used to argue for the ill-formedness of s+STOP onsets comes from Sanskrit perfect reduplication (Steriade 1984). I show, however, that the Sanskrit facts themselves do not provide any evidence for the ill-formedness of s+STOP clusters. I argue that, in the case of roots beginning with obstruent clusters, reduplication is constrained by relative markedness of fricatives and stops.

The perfect reduplicative prefix in Sanskrit consists of a CV syllable whose segmental material is entirely copied from the verbal root. When the root begins with a core cluster, the first member of the cluster is consistently copied into the reduplicant. If the root begins with an obstruent cluster, the stop is always copied into the reduplicant, regardless of whether it constitutes the first or second member of the clusters. The different behavior of the two classes of clusters is shown in the following data taken from Steriade (1982).

(31) Root Perfect Gloss

a. Core clusters:
prac\textsuperscript{h} pa-prac\textsuperscript{h} “to ask”
dru du-druv “to run”
gla: ja-gla: “to be weary”
smi si-\textipa{\textael}mi “to smile”
sru su-\textipa{\textael}ru “to flow”
mluc mu-mluc “to set”
b. Obstruent clusters:

<table>
<thead>
<tr>
<th>SF:</th>
<th>tsar</th>
<th>ta-tsar</th>
<th>“to approach stealthily”</th>
</tr>
</thead>
<tbody>
<tr>
<td>psa:</td>
<td>pa-psa:</td>
<td>“to devour”</td>
<td></td>
</tr>
<tr>
<td>kʃip</td>
<td>ci-kʃip</td>
<td>“to throw”</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FS:</th>
<th>stu</th>
<th>tu-ftu</th>
<th>“to praise”</th>
</tr>
</thead>
<tbody>
<tr>
<td>spu</td>
<td>pu-spʰu</td>
<td>“to burst”</td>
<td></td>
</tr>
<tr>
<td>skand</td>
<td>ca-skand</td>
<td>“to leap”</td>
<td></td>
</tr>
</tbody>
</table>

In her analysis of Sanskrit, Steriade interprets the reduplication patterns as evidence for the heterosyllabicity of s+STOP clusters. Since s+STOP clusters, of all the clusters, are the only ones that reduplicate the second segment of the cluster rather than the first one, she argues that they must be heterosyllabic. In particular, in her analysis, these clusters do not form regular onsets, since the initial fricative is a stray consonant, i.e. a consonant left unassociated to a syllable position. Because the initial /s/ is stray, it is invisible to reduplication and therefore cannot be copied into the reduplicant. Reduplication, thus, copies the first syllabically associated member of the cluster.

I will argue that the Sanskrit data does not provide evidence for the heterosyllabicity of s+STOP clusters. The pattern of reduplication is not based on the first syllabically associated member of the cluster, but is rather on relative sonority and relative markedness. In particular, in the case of core clusters, the least sonorous member of the cluster is reduplicated. In the case of obstruent
clusters, instead, reduplication copies the least marked obstruent, i.e. the stop\textsuperscript{13}. This proposal is based on the assumption that, unlike core clusters, obstruent clusters are not regulated by principles of sonority since fricatives and stops are assigned the same sonority value on a universal scale.

In the next section, I will first provide an analysis of the pattern of reduplication in obstruent clusters. I will only briefly discuss reduplication in the case of core clusters, since it requires a sonority-based formalism which is beyond the scope of this dissertation\textsuperscript{14}.

\subsection*{5.5.1 The Analysis of Reduplication of Obstruent Clusters}

The data showing the reduplication pattern in the case of obstruent clusters is repeated below.

\begin{equation}
\text{(32)}
\end{equation}

\text{SF: } \begin{array}{ll} 
\text{tsar} & \text{ta-\text{tsar}} \\
\text{psa:} & \text{pa-\text{psa:}} \\
\text{k} & \text{ci-k} \\
\text{ip} & \text{ip} \\
\end{array}
\begin{array}{l}
\text{“to approach stealthily”} \\
\text{“to devour”} \\
\text{“to throw”}
\end{array}

\text{FS: } \begin{array}{ll} 
\text{stu} & \text{tu-\text{stu}} \\
\text{spu} & \text{pu-sp}^{h} \text{u} \\
\text{skand} & \text{ca-skand}
\end{array}
\begin{array}{l}
\text{“to praise”} \\
\text{“to burst”} \\
\text{“to leap”}
\end{array}

\textsuperscript{13} \text{Relative sonority could be invoked also in the case of obstruent clusters under the assumption that fricatives and stops differ in sonority. In both FS and SF clusters, the stop is reduplicated, which would be the least sonorous of the two (Hironymous 1999). As argued in the previous chapters, however, stops and fricatives must crucially be equal in sonority in order to explain the phonotactics of obstruent clusters. This analysis is therefore untenable in the context of this dissertation.}

\textsuperscript{14} \text{See Clements (1989), Smolensky (1995) and Hironymous (1999) for a discussion of relative complexity of CV syllables and core clusters.}
The data clearly shows that the reduplicated morpheme consists of the [-continuant] segment of the cluster irrespective of its position in the cluster. Unlike Steriade’s characterization, reduplication is not taken here to be about first or second member of the cluster, but rather about relative markedness on the dimension of the feature [continuant]. In this respect, I argue that stops are the least marked obstruents because they are more harmonic than fricatives with respect to the relevant markedness hierarchy on the dimension of the feature [continuant]:

\[ *[+\text{continuant}] >> *[-\text{continuant}] \]

The fixed ranking in (33) is based on the fact that typologically stops are more common segments than fricatives. Moreover, as already pointed out earlier, there are languages that lack fricatives but no languages that lack stops (Maddieson 1984).

The fact that the least marked member of the cluster is copied into the reduplicant is not a surprise, since the product of reduplication is often a less marked structure than the one present in the root (McCarthy and Prince 1994). This pattern of reduplication is a type of The Emergence of the Unmarked, i.e. TETU. The idea is that the phonologically unmarked structure emerges in a certain domain, in which the more marked structure is banned, though the former is not required in the language as a whole.
My analysis of Sanskrit reduplication is based on the model of Correspondence Theory proposed in McCarthy & Prince (1995). According to this model, correspondence relations exist between the input and the output and between the base and the reduplicant.

In Sanskrit the fact that only one consonant is copied into the reduplicant is determined by the fact that the markedness constraint *COMPLEX dominates the faithfulness constraint that regulates segment correspondence between the base and the reduplicant. The constraint is defined below:

(34) MAX-BR – Every segment of the base has a correspondent in the reduplicant.

Tableau (35) below shows how the interaction of these two constraints determines the shape of the reduplicant in Sanskrit.

(35)

<table>
<thead>
<tr>
<th>Perf+tsar</th>
<th>*COMPLEX</th>
<th>MAX-BR</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. tsa-tsar</td>
<td>**!</td>
<td></td>
</tr>
<tr>
<td>b. ta-tsar</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

Both candidates (a) and (b) incur a violation of *COMPLEX because of the complex onset in the base. However, candidate (a) violates it twice because the reduplicant also contains a complex onset. Given the fact that a violation of *COMPLEX is worse than a failure to full copy, the optimal shape of the reduplicant is a simple onset. I argue that the choice of which of the two obstruents is copied into the reduplicant is determined by markedness.
Besides the fixed hierarchy in (33), the following two constraints are used in the analysis.

(36) IO-Ident(cont) – Output correspondents of an input \([\alpha\text{continuant}]\) segment are also \([\alpha\text{continuant}]\)

(37) BR-Ident(cont) – Reduplicant correspondents of a base \([\alpha\text{continuant}]\) segment are also \([\alpha\text{continuant}]\).

Both IO-Ident(cont) and BR-Ident(cont) must dominate the markedness constraint in order to predict reduplication of the \([-\text{continuant}]\) segment with a root in which both segments occur, but also allow reduplication of the \([+\text{continuant}]\) segment from a root without a complex onset. The first two tableaux below show reduplication in the case of a root beginning with a FS cluster and a SF cluster respectively. The last tableau shows, instead, reduplication in the case of a hypothetical root beginning with a fricative in a simple onset.

(38)

<table>
<thead>
<tr>
<th>Perf+stu</th>
<th>BR-Ident(cont)</th>
<th>*[+cont]</th>
<th>*[-cont]</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. (s_{ij}u-f_{ij}t_{ij}u)</td>
<td>t!</td>
<td>**</td>
<td>*</td>
</tr>
<tr>
<td>b. (t_{ij}u-f_{ij}t_{ij}u)</td>
<td>(f!)</td>
<td>*</td>
<td>**</td>
</tr>
<tr>
<td>c. (s_{ij}u-f_{ij}t_{ij}u)</td>
<td>t!</td>
<td>**</td>
<td>*</td>
</tr>
<tr>
<td>d. (t_{ij}u-f_{ij}t_{ij}u)</td>
<td>(f!)</td>
<td>*</td>
<td>**</td>
</tr>
<tr>
<td>e. (s_{ij}u-f_{ij}t_{ij}u)</td>
<td>**!</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>f. (t_{ij}u-f_{ij}t_{ij}u)</td>
<td>*</td>
<td>**</td>
<td></td>
</tr>
</tbody>
</table>
In this tableau violation of the BR-Ident(cont) has been indicated by the segment responsible for the violation. I have considered all possible correspondence relations between the base and the reduplicant. In candidates (a) and (b) both segments of the complex onset of the root have the same correspondent in the reduplicant. In both candidates (a) and (c) the violation of BR-Ident(cont) is due to the fact that the correspondent of /t/ in the reduplicant has the positive value of the feature [continuant] rather than the negative value. In candidate (b) and (d), on the contrary, the correspondent of /s/ is a [-continuant] segment. Candidate (e) and (f) satisfy BR-Ident(cont). In candidate (e) and (f), the two corresponding segments have the same values for the feature continuant. The /t/ of the root morpheme in candidate (e) and the [f] of candidate (f) have no correspondents in the reduplicant and, consequently, do not violate the constraint. Candidate (e), however, loses over candidate (f) because it receives an additional violation of *+[cont] for the fricative in the reduplicant. Candidate (f) is therefore the optimal candidate because it contains a more harmonic structure in the reduplicant, i.e. a [-cont] segment.

In tableau (39) below, which contains a root with an initial SF cluster, I have only considered the two candidates that satisfy BR-Ident(cont) as explained above.
Also in this case, the candidate with the stop in the reduplicant is the most harmonic of the two.

The constraint ranking introduced in this chapter, besides predicting the observed pattern of reduplication, allows a root [+cont] segment to be reduplicated when it occurs as a simple onset, as shown in the following tableau:

(40)

<table>
<thead>
<tr>
<th>Perf+sai</th>
<th>BR-Ident(cont)</th>
<th>*[+cont]</th>
<th>*[-cont]</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. sa-sai</td>
<td></td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>b. ta-sai</td>
<td></td>
<td>!</td>
<td></td>
</tr>
</tbody>
</table>

In the case of core clusters the reduplicated segment is selected on the basis of relative sonority, i.e. the segment lower in sonority is selected irrespective of its position in the onset, as the data in (31a) shows. Hironymous (1999) accounts for such patterns by means of a constraint that evaluates the steepest sonority cline from the edge of the syllable to the nucleus. I will informally call this constraint “Sonority Cline”\(^{15}\). In the case of core clusters, the

\(^{15}\) The constraint proposed in Hironymous (1999) is an alignment constraint that aligns consonants to the left of the syllable and ensures that such consonants will provide the steepest sonority cline from the edge of the syllable to the nucleus.
least sonorous segment is selected because it provides the steepest sonority cline, as shown in the following tableau:

(41)

<table>
<thead>
<tr>
<th>Perf+prat</th>
<th>*COMPLEX</th>
<th>SonCline</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. pa-prac</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>b. ra-prac</td>
<td>*</td>
<td>*!</td>
</tr>
<tr>
<td>c. pra-prac</td>
<td>**!</td>
<td></td>
</tr>
</tbody>
</table>

Candidate (c) fails because it contains a complex onset in the reduplicant. Both candidates (a) and (b) tie on *COMPLEX. Candidate (a) reduplicates the stop from the base and has the steepest sonority cline, thus satisfying SonCline. Candidate (b) violates SonCline because, by reduplicating the sonorant from the base, the sonority cline between onset and nucleus is minimized.

Under the assumption that fricatives and stops are assigned the same sonority value, a constraint such as SonCline would not interfere if included in the analysis of Sanskrit reduplication because both types of segments would produce the same sonority cline and thus tie on that constraint. This is shown in tableau (42) below:

(42)

<table>
<thead>
<tr>
<th>Perf+tsar</th>
<th>SonCline</th>
<th>*[+cont]</th>
<th>*[-cont]</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. sa-tsar</td>
<td>**!</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>b. ta-tsar</td>
<td>*</td>
<td>*</td>
<td>**</td>
</tr>
</tbody>
</table>
As shown, whether a fricative or a stop is reduplicated the constraint is still satisfied because both segments provide the steepest sonority cline, since as argued earlier they are equal in sonority.

In conclusion, I have shown that the Sanskrit perfect reduplication does not provide evidence for the ill-formedness of s+STOP onsets. On the contrary, it shows a pattern of reduplication in which the unmarked obstruents are copied into the reduplicant, therefore a clear case of The Emergence of the Unmarked in the reduplicant.
CONCLUSION

The main purpose of this dissertation has been to show that obstruent clusters constitute a unique phenomenon separate from core clusters. I argued that, unlike core clusters, obstruent clusters are not constrained by principles of sonority because fricatives and stops have the same sonority value. I show that a sonority-based approach is inappropriate for the phenomenon at hand.

From an empirical point of view, this dissertation contributes to the understanding of universal principles of syllable phonotactics by presenting the results of a cross-linguistic study on the occurrence and co-occurrence restrictions of obstruent clusters.

From a theoretical point of view, this dissertation contributes to the understanding and implementation of a number of tools available in Optimality Theory. In particular, I provide an explicit formalization of a technique of analysis, which I call the Subset Strategy. This strategy captures universal markedness relationships among forms without imposing any fixed ranking on the relevant constraints.

Within the proposal that obstruent clusters constitute a unique phenomenon and must be evaluated by means of principles other than sonority, this dissertation provides a new and original analysis of a long-standing problem in phonological theory, i.e. the phenomenon of s+STOP clusters. Unlike previous analyses, I argue that s+STOP clusters are special because they are unmarked
within the dimension of obstruent clusters, and not special because they are marked within the dimension of core clusters.

In what follows I summarize the content of each chapter of this dissertation and highlight its main contribution in the understanding of the phenomenon of obstruent clusters.

In Chapter 1, I provided an overview of the various issues related to sonority. I argued that obstruent clusters are different from core clusters and need therefore an analysis that does not make reference to principles of sonority. In the same chapter, I also introduced some of the optimality theoretical tools that I used in the rest of the dissertation. In particular, I introduced a method of analysis that allows us to capture markedness relationships among forms without fixed rankings.

Chapter 2, is devoted to the results that have emerged from a cross-linguistic study on the occurrence of obstruent clusters on the manner dimension. In this chapter I argued that FS clusters are the unmarked type for obstruent clusters. Moreover, I argued that the typology that results from the constraints proposed only gives rise to harmonically complete languages. Harmonically incomplete languages, i.e. the exceptions to the generalizations that I propose, are, however, found and they are argued to arise from other markedness dimensions that may interact with the proposed hierarchy.

In Chapter 3, I provided two case studies. Modern Greek exemplifies what could be called a misleading system. On the surface, a large number of obstruent
clusters are found. However, I argued that Modern Greek is in reality a fairly restrictive grammar in terms of the constraint system defining obstruent clusters. Most of the clusters found are allowed due to other independent constraints interacting with the constraints for obstruent clusters. Modern Greek also provides evidence for the unmarkedness of FS clusters because of the neutralization processes affecting SF, FF and SS clusters. The second case study is Nisqually. Nisqually provides an interesting example of a repair strategy for ill-formed obstruent clusters. In particular, in this language ill-formed clusters are repaired by obstruent syllabicity.

In Chapter 4, I have introduced the generalizations on the place dimension. In this chapter, I argued that an understanding of the phonotactics of obstruent clusters can best be understood by a system that makes explicit reference to phonetic facts. In particular, I showed that the most common pattern for obstruent clusters is coronal fricative followed by stop. I analyzed this pattern as neutralization of place in a position of weak perceptibility, i.e. the pre-consonantal position. I used English, German, Delaware and Takelma to provide examples of each language predicted by typology generated by the constraints proposed. In particular, I used Takelma as an example of a language in which the two dimensions interact in such a way as to give rise to an harmonically incomplete system, i.e. a system that violates the generalizations I propose.

Finally, in Chapter 5 I argued for the fact that s+STOP clusters are the best formed of all obstruent clusters. In particular, they are unmarked along the
manner dimension because they are FS clusters. On the place dimension, they are the least marked of all because the fricative in pre-obstruent position is a coronal. I argued against the view that the asymmetric behavior of s+STOP with respect to core clusters follows from the assumption that s+STOP clusters are marked clusters with respect to sonority. I argued, instead, that their asymmetric phonological behavior in languages such as Italian and Sanskrit follows from independent principles of syllable structure and markedness, and is not evidence for their markedness with respect to sonority.
APPENDIX I

The languages in the appendix are in alphabetical order. For each language, I indicate the references consulted and the language family to which each language belongs. Whenever possible, I have provided a full chart of obstruent clusters in onsets. These charts are intended to be representative of the native phonotactics of obstruent clusters. Marginal clusters, clusters resulting from affixation or borrowed from other languages have been either omitted from the chart or indicated with parenthesis. If relevant, such clusters have been included and discussed. Moreover, in most of the charts, voiced obstruents have been excluded unless they form clusters with voicing assimilation or mixed voicing. Parenthesis around the clusters indicate that they are rare. For some languages it has not been possible to compile full charts due to poor data sources. I have, however, included them and provided a short description of their phonotactics as in the references.
Attic Greek


Language family: Hellenic

Onset type: 5

Onset obstruent clusters:

<table>
<thead>
<tr>
<th></th>
<th>p</th>
<th>t</th>
<th>k</th>
<th>b</th>
<th>d</th>
<th>g</th>
<th>pʰ</th>
<th>tʰ</th>
<th>kʰ</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td></td>
<td>pt</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ps</td>
</tr>
<tr>
<td>t</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ks</td>
</tr>
<tr>
<td>b</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pʰ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tʰ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kʰ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>z</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Cambodian

Data source: Nacaskul, K. (1978)

Language family: Austro-Asiatic (Mon-Khmer)

Onset type: 5

### Onset Obstruent Clusters:

<table>
<thead>
<tr>
<th></th>
<th>p</th>
<th>t</th>
<th>c</th>
<th>k</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td></td>
<td>pt</td>
<td>pc</td>
<td>pk</td>
<td>ps</td>
</tr>
<tr>
<td>t</td>
<td>tp</td>
<td></td>
<td></td>
<td></td>
<td>tk</td>
</tr>
<tr>
<td>c</td>
<td>cp</td>
<td></td>
<td></td>
<td></td>
<td>ck</td>
</tr>
<tr>
<td>k</td>
<td>kp</td>
<td>kt</td>
<td>kc</td>
<td></td>
<td>ks</td>
</tr>
<tr>
<td>s</td>
<td>sp</td>
<td>st</td>
<td></td>
<td></td>
<td>sk</td>
</tr>
</tbody>
</table>

Notes: In all SS clusters, the first stop is aspirated. There are also two voiced implosives that have not been included in the chart but can form clusters.
Dakota

Data source: Boas and Deloria (1972).

Language family: Siouan (spoken throughout central and southeastern North America)

Onset type: 5

Onset Obstruent Clusters:

<table>
<thead>
<tr>
<th></th>
<th>p</th>
<th>t</th>
<th>k</th>
<th>s</th>
<th>f</th>
<th>x</th>
<th>tf</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>pt</td>
<td>ps</td>
<td>pʃ</td>
<td>ptʃ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>tk</td>
<td></td>
</tr>
<tr>
<td>k</td>
<td>kp</td>
<td>kt</td>
<td>ks</td>
<td>kʃ</td>
<td></td>
<td>ktʃ</td>
<td></td>
</tr>
<tr>
<td>s</td>
<td>sp</td>
<td>st</td>
<td>sk</td>
<td></td>
<td>stʃ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f</td>
<td>fp</td>
<td>ft</td>
<td>fk</td>
<td>fʃ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>x</td>
<td>xp</td>
<td>xt</td>
<td></td>
<td>xtʃ</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: The language also contain voiced, aspirated and glottalized series that do not form clusters.
Dutch

Data source: De Schutter (1994).

Language family: Indo-European (Germanic)

Onset type: 2

Onset Obstruent Clusters:

<table>
<thead>
<tr>
<th></th>
<th>p</th>
<th>t</th>
<th>k</th>
<th>f</th>
<th>s</th>
<th>x</th>
</tr>
</thead>
<tbody>
<tr>
<td>f</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s</td>
<td>sp</td>
<td>st</td>
<td>(sk)</td>
<td>(sf)</td>
<td>sx</td>
<td></td>
</tr>
<tr>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
English

Data source: Kenstowicz (1994)

Language family: Indo-European/Germanic

Onset type: 1

Onset Obstruent Clusters:

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>p</th>
<th>k</th>
<th>f</th>
<th>s</th>
<th>j</th>
<th>θ</th>
</tr>
</thead>
<tbody>
<tr>
<td>f</td>
<td></td>
<td></td>
<td></td>
<td>f</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s</td>
<td></td>
<td></td>
<td>st</td>
<td>sp</td>
<td>sk</td>
<td></td>
<td></td>
</tr>
<tr>
<td>j</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>θ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Georgian**

Data source: Chitoran (1994); Deprez (1988); Vogt (1971)

Language family: Caucasian

Onset type: 6

**Onset Obstruent Clusters:**

<table>
<thead>
<tr>
<th></th>
<th>b</th>
<th>d</th>
<th>g</th>
<th>pʰ</th>
<th>tʰ</th>
<th>kʰ</th>
<th>p'</th>
<th>t'</th>
<th>k'</th>
<th>q'</th>
<th>dz</th>
<th>dʒ</th>
<th>tʃʰ</th>
<th>tʃʰ</th>
<th>s</th>
<th>j</th>
<th>ɾ</th>
<th>z</th>
<th>s</th>
<th>ʒ</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>pʰ</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tʰ</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kʰ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p'</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t'</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>v</td>
<td>v</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>q'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>v</td>
<td>v</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dz</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dʒ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tʃʰ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tʃʰ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>j</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ɾ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ɾ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>z</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ʒ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
</tbody>
</table>

Notes: Harmonic clusters are indicated with x. These clusters form onsets both word initially and word medially. As noted by Chitoran, homogeneity of laryngeal features across a cluster is not necessarily associated with a clear phonetic behavior of that cluster as a complex segment, but rather with an ambiguous phonological status. For my purposes I consider them clusters rather than single segments. Harmonic clusters restrict the number of possible clusters. Checks in the chart indicate non-harmonic clusters.
**German**

Data source: Hall (1992)

Language family: Indo-European/Germanic

Onset type: 1

Onset Obstruent Clusters:

<table>
<thead>
<tr>
<th></th>
<th>ç</th>
<th>f</th>
<th>s</th>
<th>f</th>
<th>k</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>ç</td>
<td>çt</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f</td>
<td></td>
<td>(ft)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s</td>
<td></td>
<td>st</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f</td>
<td></td>
<td>ft</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k</td>
<td></td>
<td>ks</td>
<td>kt</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p</td>
<td></td>
<td></td>
<td>(pf)</td>
<td>ps</td>
<td>pt</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td></td>
<td>ts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pf</td>
<td></td>
<td>pfs</td>
<td></td>
<td>pft</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ts</td>
<td></td>
<td>tst</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tf</td>
<td></td>
<td>tft</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Parentheses indicate rare clusters. See chapter 5 for a complete analysis of German phonotactics.
**Hebrew (Modern)**

Data source: Galit Adam and Adam Ussishkin (p.c.)

Language family: Semitic (Afro-Asiatic)

Onset type: 5

**Onset Obstruent Clusters:**

<table>
<thead>
<tr>
<th></th>
<th>p</th>
<th>t</th>
<th>k</th>
<th>b</th>
<th>d</th>
<th>g</th>
<th>f</th>
<th>v</th>
<th>s</th>
<th>z</th>
<th>j</th>
<th>x</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>pt</td>
<td>pk</td>
<td>pd</td>
<td>pg</td>
<td>ps</td>
<td>pz</td>
<td>px</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>tk</td>
<td>tg</td>
<td>tf</td>
<td>ts</td>
<td>tz</td>
<td>tʃ</td>
<td>tx</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k</td>
<td>kt</td>
<td>kd</td>
<td>ks</td>
<td>kz</td>
<td>kʃ</td>
<td>kx</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>bt</td>
<td>bk</td>
<td>bd</td>
<td>bg</td>
<td>bs</td>
<td>bz</td>
<td>bx</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>dk</td>
<td>dg</td>
<td>df</td>
<td>dv</td>
<td>dʃ</td>
<td>dx</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g</td>
<td>gd</td>
<td>gf</td>
<td>gv</td>
<td>gs</td>
<td>gz</td>
<td>gʃ</td>
<td>gx</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>tv</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>v</td>
<td>(sp)</td>
<td>st</td>
<td>sk</td>
<td>sd</td>
<td>sg</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s</td>
<td>(j)</td>
<td>jk</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>z</td>
<td>zk</td>
<td>zd</td>
<td>zg</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>j</td>
<td>(ʃ)</td>
<td>ʃt</td>
<td>ʃk</td>
<td>ʃg</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Hindi

Data source: Nagamma Reddy (1987)

Language family: Indo-Aryan

Onset type: 1

Onset Obstruent Clusters:

<table>
<thead>
<tr>
<th></th>
<th>p</th>
<th>t</th>
<th>t̂</th>
<th>k</th>
<th>f</th>
<th>s</th>
<th>ŋ</th>
</tr>
</thead>
<tbody>
<tr>
<td>f</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s</td>
<td></td>
<td>sp</td>
<td>st</td>
<td></td>
<td></td>
<td></td>
<td>sk</td>
</tr>
<tr>
<td>ŋ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: The language also contains voiced and aspirated stops that do not occur in clusters.
Italian


Language family: Indo-European (Romance)

Onset type: 1

Onset Obstruent Clusters:

<table>
<thead>
<tr>
<th></th>
<th>p</th>
<th>t</th>
<th>k</th>
<th>f</th>
<th>s</th>
<th>j</th>
</tr>
</thead>
<tbody>
<tr>
<td>f</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s</td>
<td></td>
<td>sp</td>
<td>st</td>
<td>sk</td>
<td></td>
<td>(sf)</td>
</tr>
<tr>
<td>j</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Italian has voicing assimilation in obstruent clusters. Therefore the clusters [zb zd zg] are all attested. [sf] clusters are only found in morphologically complex words.
## Khasi

Data source: Henderson (1976)

Language family: Mon-Khmer

Onset type: 5

### Onset Obstruent Clusters:

<table>
<thead>
<tr>
<th></th>
<th>p</th>
<th>t</th>
<th>k</th>
<th>pʰ</th>
<th>tʰ</th>
<th>kʰ</th>
<th>b</th>
<th>d</th>
<th>s</th>
<th>f</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>pt</td>
<td></td>
<td></td>
<td>pd</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>tk</td>
<td></td>
<td></td>
<td>tb</td>
<td>td</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k</td>
<td>kp</td>
<td>kt</td>
<td>kʰ</td>
<td>kb</td>
<td>kd</td>
<td>ks</td>
<td>kʃ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pʰ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tʰ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kʰ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b</td>
<td></td>
<td>bt</td>
<td></td>
<td>bd</td>
<td>bs</td>
<td>bʃ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>dp</td>
<td>dk</td>
<td>dkʰ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s</td>
<td>sp</td>
<td>st</td>
<td>sk</td>
<td></td>
<td>sb</td>
<td>s d</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f</td>
<td>f t</td>
<td>f k</td>
<td></td>
<td>f d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# Lithuanian

Data source: Tankeviciute and Strimaitiene (1990)

Language family: Indo-European (Baltic)

Onset type: 1

## Onset Obstruent Clusters:

<table>
<thead>
<tr>
<th></th>
<th>p</th>
<th>t</th>
<th>k</th>
<th>s</th>
<th>ʃ</th>
</tr>
</thead>
<tbody>
<tr>
<td>s</td>
<td>sp</td>
<td>st</td>
<td>sk</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ʃ</td>
<td>ʃp</td>
<td>ʃt</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: The voiced fricative /v/ also occurs in clusters. It however seems to pattern with the sonorants as in a number of other languages.
Isthmus Zapotec

Data source: Marlett and Pickett (1987)

Language family: Zapotec/Amerindian

Onset type: 1

Onset Obstruent Clusters:

<table>
<thead>
<tr>
<th></th>
<th>p</th>
<th>t</th>
<th>k</th>
<th>s</th>
<th>f</th>
<th>tf</th>
</tr>
</thead>
<tbody>
<tr>
<td>s</td>
<td>st</td>
<td>sk</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f</td>
<td>fp</td>
<td>ft</td>
<td>fk</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: These clusters are mostly found in morphologically complex words (possessed forms of nouns). Although, rarely they are also found in monomorphemic words. For this reason I am assuming that FS clusters are indeed well-formed in the language.
Haida

Data source: Swanton (1910); Lawrence (1977).

Language family: Isolate

Onset type: 1

Onset Obstruent Clusters:

<table>
<thead>
<tr>
<th></th>
<th>p</th>
<th>t</th>
<th>k</th>
<th>q</th>
<th>ʔq</th>
<th>ts</th>
<th>ʔt</th>
</tr>
</thead>
<tbody>
<tr>
<td>s</td>
<td>sp</td>
<td>st</td>
<td>sk</td>
<td>sq</td>
<td>ʔq</td>
<td>stʔ</td>
<td></td>
</tr>
<tr>
<td>ṭ</td>
<td>ṭp</td>
<td>ṭt</td>
<td>ṭk</td>
<td>ṭq</td>
<td>ʔq</td>
<td></td>
<td></td>
</tr>
<tr>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ʔx</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: The language also contains a glottalized and an aspirated series of stops and affricates. Both series occur in FS clusters.
Mawo

Data source: Hongkai (1986); Namkung (1996).

Language family: Qiang (Tibeto-Burman)

Onset type: 3

Onset Obstruent Clusters:

\[
\begin{array}{c|cccccccccccc}
\text{p} & \text{t} & \text{ts} & \text{t§} & \text{t¢} & \text{k} & \text{q} & \text{s} & \text{§} & \text{¢} & \text{x} & \text{χ} \\
\hline
\text{k}^b & \text{k}^b s & \text{k}^b § & \text{k}^b ¢ & \\
\text{q}^b & \text{q}^b s & \text{q}^b § & \\
\text{s} & \text{sp} & \text{st} & \text{st¢} & \text{sk} & \text{sq} & \\
\text{§} & \text{sp} & \text{st¢} & \text{sk} & \text{sq} & \\
\text{¢} & \text{xp} & \text{xts} & \text{xt§} & \text{xt¢} & \\
\text{x} & \text{xt} & \text{xts} & \text{xt§} & \text{xt¢} & \\
\text{χ} & \text{χt} & \text{χts} & \text{χt§} & \text{χt¢} & \\
\end{array}
\]

Notes: A large number of the onset clusters arise from affixation. Many affixes in Mawo consist of a single fricative.
### Misantla Totonac


Language family: language isolate

Onset type: 1

**Onset Obstruent Clusters:**

<table>
<thead>
<tr>
<th></th>
<th>p</th>
<th>t</th>
<th>k</th>
<th>q</th>
<th>s</th>
<th>l</th>
<th>ts</th>
<th>tsf</th>
</tr>
</thead>
<tbody>
<tr>
<td>s</td>
<td>sp</td>
<td>st</td>
<td>sk</td>
<td>sq</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>l</td>
<td>lp</td>
<td>lt</td>
<td>lk</td>
<td>lq</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Nisgha

Data source: Tarpent (1989)

Language family: Tsimshianic

Onset type: 6

Onset Obstruent Clusters:

<table>
<thead>
<tr>
<th></th>
<th>p</th>
<th>t</th>
<th>k</th>
<th>k^w</th>
<th>q</th>
<th>s</th>
<th>ḩ</th>
<th>x</th>
<th>x^w</th>
<th>χ</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>pt</td>
<td>pc</td>
<td></td>
<td></td>
<td>px</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td></td>
<td>tk</td>
<td>tk^w</td>
<td></td>
<td>tx</td>
<td>tχ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c</td>
<td></td>
<td></td>
<td>ks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k</td>
<td></td>
<td></td>
<td></td>
<td>k^w</td>
<td></td>
<td>k^w</td>
<td>k^w</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>q</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s</td>
<td>sp</td>
<td>st</td>
<td>sk</td>
<td>sk^w</td>
<td>sq</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ḩ</td>
<td>ḩp</td>
<td>ḩk</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>x</td>
<td>xp</td>
<td>xt</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>x^w</td>
<td></td>
<td>x^w</td>
<td></td>
<td>x^w</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>χ</td>
<td>χp</td>
<td>χ</td>
<td></td>
<td>χq</td>
<td>χs</td>
<td>χ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: These are only representative. Other clusters occur that contain also the glottalized variants. The cluster [χs] was found only in one word, I will consider it marginal. SS clusters whose final member is either [k^w] or [t] are mostly the result of affixation. According to my source, the language is mostly characterized by an alternance of S and F in either order with some exceptions.
Pashto

Data source: Penzl (1995)

Language family: Indo-Iranian

Onset type: 4

Onset Obstruent Clusters:

<table>
<thead>
<tr>
<th></th>
<th>p</th>
<th>t</th>
<th>ť</th>
<th>k</th>
<th>s</th>
<th>j</th>
<th>Ꞅ</th>
<th>x</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ps</td>
<td>ps</td>
<td>ps</td>
<td></td>
</tr>
<tr>
<td>t</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ts</td>
<td>ts</td>
<td>ts</td>
<td></td>
</tr>
<tr>
<td>ť</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ǩ</td>
<td></td>
</tr>
<tr>
<td>k</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ǩ</td>
</tr>
<tr>
<td>s</td>
<td>sp</td>
<td>st</td>
<td>sť</td>
<td>sk</td>
<td></td>
<td></td>
<td></td>
<td>sx</td>
</tr>
<tr>
<td>ǰ</td>
<td>ǰp</td>
<td>ǰt</td>
<td>ǰť</td>
<td>ǰk</td>
<td></td>
<td></td>
<td>ǰx</td>
<td></td>
</tr>
<tr>
<td>Ꞅ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ꞅk</td>
<td>Ꞅk</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x̌</td>
<td></td>
</tr>
</tbody>
</table>

Note: Pashto has voicing assimilation in obstruent clusters. All obstruents have voiceless and voiced phonemes.
Serbo-Croatian

Data source: Hodge (1946)

Language family: Indo-European/Slavic

Onset type: 6

Onset Obstruent Clusters:

<table>
<thead>
<tr>
<th></th>
<th>p</th>
<th>t</th>
<th>k</th>
<th>f</th>
<th>s</th>
<th>ŋ</th>
<th>ċ</th>
<th>ć</th>
<th>b</th>
<th>d</th>
<th>g</th>
<th>ď</th>
<th>ţ</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>pt</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td></td>
<td>tk</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ŋ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ċ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ć</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ď</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ţ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: [c ć ć'] indicate respectively dental, alveolar and post-alveolar voiceless affricates. [ď] represents a voiced alveolar affricate. The onset clusters also includes recent loans. [v] has not been included since it patterned with sonorants.
Seri

Data source: Marlett (1988)

Language family: Hokan

Onset type: 6

Onset Obstruent Clusters:

<table>
<thead>
<tr>
<th></th>
<th>p</th>
<th>t</th>
<th>k</th>
<th>ɸ</th>
<th>s</th>
<th>ɬ</th>
<th>ɬ</th>
<th>x</th>
<th>ɣ</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>pt</td>
<td></td>
<td></td>
<td></td>
<td>ps</td>
<td></td>
<td></td>
<td></td>
<td>pɣ</td>
</tr>
<tr>
<td>t</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>tɬ</td>
<td>tx</td>
<td></td>
<td></td>
</tr>
<tr>
<td>k</td>
<td>kp</td>
<td>kt</td>
<td></td>
<td></td>
<td></td>
<td>ks</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ɸ</td>
<td>sp</td>
<td>st</td>
<td>sk</td>
<td></td>
<td>sl</td>
<td></td>
<td></td>
<td></td>
<td>ɬɣ</td>
</tr>
<tr>
<td>s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>xk</td>
<td></td>
</tr>
<tr>
<td>ɬ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ɣx</td>
</tr>
</tbody>
</table>
Telugu

Data source: Nagamma Reddy (1987)

Language family: Dravidian

Onset type: 1

Onset Obstruent Clusters:

<table>
<thead>
<tr>
<th></th>
<th>p</th>
<th>t</th>
<th>t</th>
<th>k</th>
<th>f</th>
<th>s</th>
<th>š</th>
</tr>
</thead>
<tbody>
<tr>
<td>f′</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s</td>
<td></td>
<td>sp</td>
<td>st</td>
<td></td>
<td></td>
<td>sk</td>
<td></td>
</tr>
<tr>
<td>Š</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Š</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: The language also contains voiced and aspirated stops that do not occur in clusters.
**Tsou**

Data source: Wright (1996)

Language family: Austronesian

Onset type: 6

Onset Obstruent Clusters:

<table>
<thead>
<tr>
<th></th>
<th>p</th>
<th>t</th>
<th>k</th>
<th>ʔ</th>
<th>b</th>
<th>d</th>
<th>f</th>
<th>v</th>
<th>s</th>
<th>z</th>
<th>ts</th>
<th>h</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>pt</td>
<td>pk</td>
<td>pʔ</td>
<td>pd</td>
<td>ps</td>
<td>pts</td>
<td>ph</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>tp</td>
<td>tk</td>
<td>tʔ</td>
<td>tɓ</td>
<td>tf</td>
<td>tv</td>
<td>th</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k</td>
<td>ʔp</td>
<td>ʔt</td>
<td>ʔv</td>
<td>ʔs</td>
<td>ʔh</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ɓ</td>
<td>ɓs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d’</td>
<td>f’t</td>
<td>fk</td>
<td>fʔ</td>
<td>f’s</td>
<td>fzs</td>
<td>fts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>v</td>
<td>vz</td>
<td>vts</td>
<td>vh</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s</td>
<td>sp</td>
<td>sk</td>
<td>sʔ</td>
<td>sɓ</td>
<td>sf</td>
<td>sv</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>z</td>
<td>zv</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ts</td>
<td>tsp</td>
<td>tsk</td>
<td>tsʔ</td>
<td>tsf</td>
<td>tsv</td>
<td>tsz</td>
<td>tsh</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Wichita

Data source: Rood (1975)

Language family: Caddoan

Onset type: 3

Onset Obstruent Clusters:

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>k</th>
<th>k^w</th>
<th>ts</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k</td>
<td></td>
<td></td>
<td></td>
<td>ks</td>
<td></td>
</tr>
<tr>
<td>k^w</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ts</td>
<td></td>
<td></td>
<td></td>
<td>tsk</td>
<td></td>
</tr>
<tr>
<td>s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>sk</td>
</tr>
</tbody>
</table>

Notes: Wichita represents a very unusual system. The obstruent system is very restricted. This results in a very restricted set of onset clusters as well. The affricate /ts/ combines with a following /k/. This may indicate that /ts/ patterns with /s/ rather than stops since no SS clusters are found in the language. The three consonantal cluster /ksk/ is also found in monomorphemic words.
**Yatee Zapotec**

Data source: Jager and Van Valin (1982)

Language family: Zapotecan

Onset type: 5/6 depending on the status of FF clusters.

<table>
<thead>
<tr>
<th>Onset Obstruent Clusters:</th>
</tr>
</thead>
<tbody>
<tr>
<td>p t k b d g s z š z̆ tʃ dʒ</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>p</th>
<th>pʰ</th>
<th>pʃ</th>
<th>pʃ̆</th>
<th>pʃ̆̆</th>
<th>pʃ̆̆̆</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>tʃ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>bd</td>
<td>bg</td>
<td>bz</td>
<td>bʒ</td>
<td>bʒ̆</td>
</tr>
<tr>
<td>d</td>
<td></td>
<td></td>
<td></td>
<td>dz</td>
<td>dz̆</td>
</tr>
<tr>
<td>s</td>
<td>st</td>
<td></td>
<td></td>
<td>stʃ</td>
<td></td>
</tr>
<tr>
<td>z</td>
<td>zb</td>
<td>zd</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>š</td>
<td>št</td>
<td>šk</td>
<td>šs</td>
<td>štʃ</td>
<td></td>
</tr>
<tr>
<td>z̆</td>
<td>z̆b</td>
<td>z̆d</td>
<td>zg</td>
<td>z̆s</td>
<td>z̆dʒ</td>
</tr>
<tr>
<td>tʃ</td>
<td></td>
<td></td>
<td></td>
<td>tʃ</td>
<td></td>
</tr>
<tr>
<td>dʒ</td>
<td></td>
<td></td>
<td></td>
<td>dʒ</td>
<td>dʒ̆</td>
</tr>
<tr>
<td>ų</td>
<td>ųt</td>
<td></td>
<td></td>
<td>ųʃ</td>
<td>ųdʒ</td>
</tr>
</tbody>
</table>

Notes: /ʒ/ surfaces as [ʃ] in the environment of voiceless consonants. Many of the sibilant clusters arise from affixation of the continuative aspect prefixes [š-] and [z̆-].
Yuchi

Data source: Wolff (1948); Crawford (1973).

Language family: language isolate (maybe Siouan)

Onset type: 1

Onset Obstruent Clusters:

<table>
<thead>
<tr>
<th></th>
<th>p</th>
<th>t</th>
<th>k</th>
<th>f</th>
<th>s</th>
<th>ʃ</th>
<th>ɬ</th>
<th>ts</th>
<th>tʃ</th>
</tr>
</thead>
<tbody>
<tr>
<td>f</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ʃp</td>
<td>ʃt</td>
<td>ʃk</td>
<td></td>
</tr>
<tr>
<td>s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ʃp</td>
<td>ʃt</td>
<td>ʃk</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ɬ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ʃp</td>
<td>ʃt</td>
<td>ʃk</td>
</tr>
</tbody>
</table>

Notes: The language contains also voiced, aspirated and glottalized stops, as well as glottalized fricatives. Except for /t'/ and /k'/ no other stop is found in clusters. Affricates are not allowed in clusters.
Other languages in the typology for which it was not possible to construct a chart:

<table>
<thead>
<tr>
<th>Language</th>
<th>FS</th>
<th>SF</th>
<th>SS</th>
<th>FF</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eggon</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>Maddieson (1981)</td>
</tr>
<tr>
<td>Havasupai</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>Kozlowski (1976)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Seiden (1963)</td>
</tr>
<tr>
<td>Nisqually</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>Hoard (1978)</td>
</tr>
</tbody>
</table>


