INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

UNIVERSITY OF CALIFORNIA

Los Angeles

Laryngeal Cooccurrence Restrictions

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Linguistics

by

Margaret R. MacEachern

1997

The dissertation of Margaret R. MacEachern is approved.

Bruce Hayes

Donka Minkova

Donca Steriade, Committee Chair

University of California, Los Angeles

1997

TABLE OF CONTENTS

LIST OF TABLES

ACKNOWLEDGMENTS

I thank my committee members for their guidance and encouragement: Donka Minkova, Ed Stabler, and in particular Bruce Hayes and Donca Steriade. I am grateful for the many, many hours they have spent teaching me linguistics over the past five years.

I also want to thank all my friends in the Department of Linguistics. I found in preparing the list that there were too many people to mention them all here individually, but a complete list would include Victoria Anderson, Filippo Beghelli, Katherine Crosswhite, Edward Flemming, Ed Garrett, Alhaji Gimba, Matt Gordon, Chai-Shune Hsu, Jongho Jun, Abby Kaun, Heather Morrison, Kie Ross, Bonny Sands, Michael Shalev, Dan Silverman, Kim Thomas, Richard Wright, and Jie Zhang.

I had the good fortune to enter the program with a great group of people: Jim Borad, Manuel Español-Echevarria, Christina Foreman, Robert Kirchner, Felicia Lee, Benjamin Munson, Brian Potter, and especially Rod Casali. Thanks again.

On the non-linguistic side of things, I am grateful to Ellen Casali and Suzanne Gross for their friendship and hospitality. They made Los Angeles a much nicer place to be.

But most of all, thanks to my family, and to Pete Morcos, for everything.

This research was supported in part by a J.K. Javits Fellowship and a University of California Dissertation Year Fellowship.

VITA

	Born, Minneapolis, Minnesota
1983-1987	B.A., *magna cum laude*, Anthropology and Sociology Carleton College Northfield, Minnesota
1991-1992	M.A., Linguistics University of Washington Seattle, Washington
1992-1996	J.K. Javits Fellowship UCLA
Summer 1993	Graduate Student Researcher for Profs. Ian Maddieson and Patricia Keating Department of Linguistics UCLA
Summer 1993	Title VI Fellowship (Quechua language study) UCLA
Fall 1994	Teaching Assistant for Prof. Sun-Ah Jun in Introduction to Phonetics Department of Linguistics UCLA
Summer 1995	Graduate student researcher for Dr. Sharon Oviatt Project on Human-Computer Interaction and Multimodal Systems Oregon Graduate Institute of Science and Technology
Fall 1995	Teaching Assistant for Prof. Anna Szabolcsi in Introduction to Linguistics Department of Linguistics UCLA
1996-1997	Dissertation Year Fellowship UCLA
Spring 1996	Teaching Assistant for Prof. Donca Steriade in Phonology I Department of Linguistics UCLA

Summer 1996 Graduate Student Researcher for Prof. Peter Ladefoged
Project on Endangered Languages
Department of Linguistics
UCLA

Summer 1996 Graduate Student Researcher for Dr. Lynne Bernstein
Spoken Language Processes and Devices Laboratory
House Ear Institute
Los Angeles, California

PUBLICATIONS AND PRESENTATIONS

Hayes, Bruce, and Margaret MacEachern, 1996. "Are there lines in folk poetry?" *UCLA Working Papers in Phonology*:125-142.

Hayes, Bruce, and Margaret MacEachern, forthcoming. "Folk verse form in English." *Language*.

Keating, Patricia A., Peggy MacEachern, Aaron Shryock, and Sylvia Dominguez, 1994. "A manual for phonetic transcription: segmentation and labeling of words in spontaneous speech." *UCLA Working Papers in Phonetics* 88:91-121.

MacEachern, Margaret, 1996. "Computer modeling of the production of ejectives using aerodynamic and acoustic data." Acoustical Society of America, Honolulu.

MacEachern, Margaret, 1996. "Ordering restrictions on aspirated and ejective stops in Bolivian Aymara." *Proceedings of the Twenty-Second Annual Meeting of the Berkeley Linguistics Society*, edited by Jan Johnson, Matthew L. Juge, and Jeri L. Moxley, pp 244-254. Berkeley, California: Berkeley Linguistics Society.

MacEachern, Margaret, 1996. "Ordering restrictions on aspirated and ejective stops in Bolivian Aymara." Southwest Optimality Theory Workshop, Irvine, California.

MacEachern, Margaret, 1997. "Laryngeal similarity effects in Quechua and Aymara." *The Proceedings of the Fifteenth West Coast Conference on Formal Linguistics*, edited by Brian Agbayani and Sze-Wing Tang, pp 337-351. Center for the Study of Language and Information, Stanford University.

MacEachern, Margaret, 1997. "Wari' Phonetic Structures." *Journal of Amazonian Languages* 1:3-28.

MacEachern, Margaret, and Peter Ladefoged, 1997. "Wari' vowels: an instrumental study of a remarkable inventory." Linguistic Society of America, Chicago.

Oviatt, Sharon, Gina-Anne Levow, Margaret MacEachern, and Karen Kuhn, 1996. "Modeling hyperarticulate speech during human-computer error resolution." *Proceedings of the Fourth International Conference on Spoken Language Processing*.

ABSTRACT OF THE DISSERTATION

Laryngeal Cooccurrence Restrictions

by

Margaret R. MacEachern

Doctor of Philosophy in Linguistics

University of California, Los Angeles, 1997

Professor Donca Steriade, Chair

Cooccurrence restrictions prohibit the coexistence of certain segments within some domain, typically the morpheme or the word. In this study, I investigate laryngeal cooccurrence restrictions; these restrictions operate over voiceless aspirated segments, voiced aspirated segments, ejectives, implosives, /h/, and /ʔ/.

I provide a typological statement of laryngeal cooccurrence restrictions, based on a study of dictionaries from eleven languages (Cuzco Quechua, Souletin Basque, Sanskrit, Aymara dialects from Peru and Bolivia, Ofo, Gojri, Old Georgian, Hausa, Tzutujil, and Shuswap). The typological statement is quite restrictive: all of the languages studied fit into four basic patterns, and these patterns are distinguished by three striking characteristics.

First, the restrictions fit a crosslinguistic implicational hierarchy: the presence or absence of a particular restriction within a language can be predicted, given the other restrictions of the language. Second, this hierarchy appears to be based on similarity: the different patterns of cooccurrence restrictions result from the point of acceptable similarity being set differently in the different languages under consideration. Finally, identical segments act in a monolithic fashion with respect to laryngeal cooccurrence restrictions: if a language allows identical aspirated stops to cooccur, for example, then (inventory factors permitting) identical ejective stops may also cooccur.

I offer an Optimality Theory analysis of laryngeal cooccurrence restrictions. In this analysis, the restrictions emerge from the interactions of a constraint that penalizes (auditory) similarity, a constraint that requires consonantal identity, and constraints that enforce faithfulness between laryngeal features in input and output forms. I also show that traditional autosegmental accounts based on the Obligatory Contour Principle, although successful in analyzing data from some languages, cannot be extended to cover everything in the typological statement.

CHAPTER 1: INTRODUCTION

Cooccurrence restrictions prohibit the coexistence of certain segments within some domain, typically the morpheme or the word. Crosslinguistically, cooccurrence restrictions vary in several respects. They differ according to the following four points:

(i) The type of segments upon which the restriction operates

(ii) The domain within which the restriction operates

(iii) The strength of the restriction – the number of counterexamples to it that exist

(iv) The "location" or scope of the restriction within the lexicon – whether it applies uniformly throughout the lexicon, or is limited to a subset of the lexicon (for example, verbal roots, or the native vocabulary)

Some examples of cooccurrence restrictions are given in Table 1.1 (see sources for details, particularly on how "homorganic" represents a simplification of the data).

Language	Restriction	Sources
Sanskrit	Roots and reduplicative constructions do not contain more than one aspirated segment: *tʰutʰ, *gʱutʰ.	Grassmann 1863, Whitney 1941, Whitney 1945, Kiparsky 1965.
Arabic	Verbal roots do not contain identical consonants; homorganic consonants are disfavored: *mmd.[1]	Greenberg 1950, McCarthy 1981, 1986, Pierrehumbert 1993, Frisch, Broe and Pierrehumbert 1995, Frisch 1996.
Javanese	Homorganic consonants do not cooccur root-internally, unless they are identical: *bamaʔ, but babot 'carpet'.	Uhlenbeck 1950, Mester 1986.
Chumash	Words do not contain sibilants that differ in place of articulation: osos 'heel', pʃoʃ 'gopher snake', but *osoʃ.	Beeler 1970, Yip 1988.
Japanese	Native roots do not include more than one voiced obstruent: *gaze.	Lyman 1894, Ito and Mester 1986.

Table 1.1. Sample cooccurrence restrictions.

The cooccurrence restrictions listed in Table 1.1 restrict several different types of segments from cooccurring: aspirated segments, identical segments, homorganic segments, non-identical sibilants, and voiced obstruents. These restrictions also operate over different levels: the Arabic, Javanese, and Japanese restrictions are root-level, the Chumash restrictions are word-level, and the Sanskrit restrictions apply to both roots and reduplicated forms (but not over roots and suffixes, for example).

[1] Arabic has a root-and-pattern morphology with consonantal roots; see McCarthy 1981 for an explanation.

The Arabic restrictions illustrate the strength effect. In that language, identical consonants are prohibited from cooccurring, while homorganic consonants do cooccur, although at much less than the expected rates. The restriction on identical segments is absolute; the restriction on homorganic elements is gradient (Greenberg 1950, McCarthy 1981, 1986, Pierrehumbert 1993, Frisch, Broe and Pierrehumbert 1995, Frisch 1996).

Japanese provides an example of a cooccurrence restriction that does not apply throughout the lexicon. In that language, two voiced obstruents are banned only in native, or "Yamato" morphemes; they are present in foreign loans and pejorative forms (buzaa 'buzzer', debu 'Fatty'; Ito and Mester 1986).

Analyses of cooccurrence restrictions have had a major impact on phonological theory in two areas. First, the long-distance nature of these constraints – the fact that they operate over several intervening segments – makes them prime objects for research on the degree to which phonology operates over abstract representations. Cooccurrence restrictions are frequently viewed as evidence for autosegmentalism and the Obligatory Contour Principle (OCP; McCarthy 1986, 1988, Mester 1986, and Yip 1989, among others). The OCP is said to operate over adjacent, identical features; features of segments that are not string-adjacent can only be adjacent under some abstract definition of adjacency.

Second, the appearance of cooccurrence restrictions in a variety of genetically diverse, geographically-dispersed languages suggests that the process is relevant to the search for phonological universals. Such data is also important to historical linguistics –

arguments for the separate development of Grassmann's Law in Sanskrit and Greek have been buttressed by examples of similar phenomena in unrelated languages such as Ofo and Salish (Thompson and Thompson 1985; see also de Reuse 1981).

In this study, I examine laryngeal cooccurrence restrictions. These restrictions operate over the following set of segments: voiceless aspirated segments, voiced aspirated segments, ejectives, implosives, /h/, and /ʔ/. The Sanskrit restriction noted in Table 1.1 is perhaps the most well-known example of a laryngeal cooccurrence restriction.

One significant result reported here is that the attested set of laryngeal cooccurrence restrictions fit into a few basic patterns: the typological statement is quite restrictive. In this research, I will suggest that the existing restrictions reflect prohibitions on the combination within morphemes (or roots) of segments that are extremely similar, in auditory terms. The different degrees of cooccurrence restrictions result from the point of acceptable similarity being set differently in the different languages under consideration.

Four patterns of laryngeal cooccurrence restrictions are schematized in Table 1.2, which records the restrictions found in one language of each type.

	Pattern One: Cuzco Quechua	Pattern Two: Peruvian Aymara[2]	Pattern Three: Bolivian Aymara	Pattern Four: Tzutujil
t^h ⇔ ʔ	✓	✓	✓	
t' ⇔ h	✓	✓	✓	✓
ɗ ⇔ h				✓
h ⇔ ʔ	✓	✓	✓	✓
t^h ⇔ h	*	✓	✓	
t' ⇔ k^h	*	*	✓	
t^h ⇔ k^h	*	*	✓	
ʔ ⇔ ɗ				✓
ʔ ⇔ t'	*	*	*	✓
t' ⇔ ɓ				✓
ɓ ⇔ ɗ				
t' ⇔ k'	*	*	*	*
t' ⇔ t^h	*	*	*	
h ⇔ h	*	✓	✓	✓
ʔ ⇔ ʔ				✓
t^h ⇔ t^h	*	✓	✓	
t' ⇔ t'	*	✓	✓	✓
ɗ ⇔ ɗ				

Table 1.2. Laryngeal cooccurrence restrictions: four patterns.

In Table 1.2, the leftmost column lists pairs of elements; it is the cooccurrence of these elements that is under consideration in the last four columns of the table. The symbols **t^h** and **k^h** designate aspirated stops; **t'** and **k'** designate ejective stops, and ɓ and ɗ designate implosives. These symbols stand for segments at any place of articulation, except when elements from the two sets {**t^h, t', ɗ**} and {**k^h, k', ɓ**} mix, or when identical

[2] I use "Peruvian Aymara" and "Bolivian Aymara" as shorthand designations of two Aymara dialects that are spoken in Peru and Bolivia, respectively. More information on these dialects is provided in 3.2.1 and 3.3.1.

elements are listed (for example, t' ⇔ t'). In the first case, the segments should be understood to be heterorganic: for example, the row labeled t^h ⇔ k^h is concerned with the cooccurrence of heterorganic aspirated stops – all forms such as t^{h}arp^{h}a, p^{h}ut^{h}u, t^{h}ilk^{h}a, etc. When identical elements are listed (as in the last five rows of Table 1.2), it is the cooccurrence of identical elements of the specified type that is under consideration – for example, forms with two /h/s (hilha, huha, etc.), forms with identical aspirated stops (t^{h}alt^{h}a, p^{h}up^{h}u, etc.), and so on.

The body of the table records the presence or absence of each type of form in one sample language for each of the four patterns under discussion. A checkmark indicates that these forms are found in the language, and there is no cooccurrence restriction against them. An asterisk indicates that such forms are absent. Shaded cells indicate that such forms are excluded due to inventory considerations (for example, the lack of implosives in Cuzco Quechua, Peruvian Aymara, and Bolivian Aymara leads to empty cells in the rows labeled ʔ ⇔ ɗ, t' ⇔ ɓ, etc.) or distributional restrictions (for example, glottal stop is found only in initial position in Cuzco Quechua, Peruvian Aymara, and Bolivian Aymara; therefore, the cells in the row labeled ʔ ⇔ ʔ are shaded for these columns).

When the data is organized as in Table 1.2, three characteristics of the typology become conspicuous. First, the data fits an implicational hierarchy: if we know the location of an asterisk or a checkmark (representing the presence and absence, respectively, of cooccurrence restrictions), then I claim that we can predict the presence

or absence of other cooccurrence restrictions in the language under consideration. For example, Table 1.2 tells us that heterorganic aspirated and ejective stops can cooccur in Bolivian Aymara morphemes, but that they do not cooccur in Cuzco Quechua or Peruvian Aymara morphemes. With just this knowledge, we could use Table 1.2 to predict – among other things – that forms containing a glottal stop and an ejective will be absent in Cuzco Quechua and Peruvian Aymara, while forms with /h/ and an aspirated stop will be present in Bolivian Aymara.

Second, Table 1.2 illustrates an interesting point concerning the cooccurrence of identical segments. When languages allow pairs of /h/s, or identical aspirated stops, etc., to cooccur, then *all* such pairs of identical segments can cooccur. For example, if we know that forms with two /h/s are found in Peruvian Aymara, Bolivian Aymara, and Tzutujil, but not in Cuzco Quechua, then we know that forms with identical ejectives will also be found in Peruvian Aymara, Bolivian Aymara, and Tzutujil, while such forms will not be found in Cuzco Quechua. In other words, identical segments act in a monolithic fashion with respect to laryngeal cooccurrence restrictions. (This description represents a simplication of the data which is redressed in chapters 4 and 5.)

The third typological observation I make on the basis of Table 1.2 is not as apparent as the first two. I suggest here that the hierarchy of elements shown in Table 1.2 is based upon auditory similarity: elements within rows are dissimilar at the top of the hierarchy, and become progressively more similar as we descend the hierarchy. Cooccurrence restrictions always prohibit a block of similar elements from cooccurring,

but the point of acceptable similarity is set differently in the different languages under consideration. This means that in a chart like Table 1.2, the restrictions will always extend up from the bottom of the table in a solid block.

Consider the similarity continuum shown in (1). Ch represents an aspirated stop at any place of articulation; C' represents an ejective at any place of articulation.

(1)

h------Ch------C'------ʔ

The nature of the continuum is as follows: /h/ and Ch have aspiration in common; Ch and C' share a lag between oral release and the onset of modal voicing in a following vowel (or temporal sequencing relations, if these consonants are not followed by vowels); C' and /ʔ/ share the property of laryngealization.

In this study, I provide an Optimality Theoretic analysis that accounts for the existence of laryngeal cooccurrence restrictions. I make the following three claims:

1. Cooccurrence restrictions fit an implicational hierarchy. In other words, the pairs of elements that are prohibited from cooccurring form a block at the top of the similarity scale. It is only the size of the block, not its placement, that varies by language.

In a Pattern One language like Cuzco Quechua, we do not find elements that are adjacent in (1) cooccurring within roots. We also do not find any of the elements in (1)

coexisting with identical copies of themselves (*halha, *t'ult'u, etc.). However, elements in (1) that are separated by at least one other element *do* cooccur: hutk'u 'tube-like hole', ʔantʃʰa 'very'. It seems that similar elements are prohibited from cooccurring, while dissimilar elements are allowed to cooccur. I discuss this notion further in chapter 4.

2. In some languages, identical elements may cooccur, even though similar – but non-identical – elements do not. There are two types of identity effects – the complete, or monolithic effect, and the partial, or incomplete identity effect (not described above; this effect is characteristic of Shuswap, and also – I will argue – of Ofo). The analysis offered here accounts for both types of identity effect.

3. Similar elements are prohibited from cooccurring. The point at which similarity becomes unacceptable is set differently in different languages, and may derive from a scale of auditory similarity, which I presume would follow from the human perceptual apparatus. In at least one respect (voice onset time), the similarity of two elements can only be characterized with respect to auditory facts – it cannot be represented in articulatory terms.[3]

[3] For important research on a different set of similarity effects in phonology, see Kaun 1995 on rounding harmony in vowel systems.

This dissertation represents only the initial steps in the exploration of laryngeal cooccurrence restrictions. Important issues for study that I have not yet addressed include the following:

1. Establishing the type of similarity at issue. I suggest here that it is auditory similarity that is relevant, but I do not have experimental confirmation of the relative auditory similarities of the segments under discussion. Other researchers have suggested that it is feature-based or natural classes-based similarity that is at issue.

2. The existence of gradient cooccurrence restrictions. Various scholars have noted that the absolute prohibitions of some languages are mirrored by gradient effects in other languages, and some recent research has addressed gradient cooccurrence restriction data (Pierrehumbert 1993, Berkley 1994, Frisch, Broe and Pierrehumbert 1995, Frisch 1996; all of this research deals with cooccurrence restrictions operating over place of articulation).

I also do not explore the lexical overrepresentation of dissimilar items. Frisch, Broe and Pierrehumbert 1995 suggest that patterns of overrepresentation of root-internal consonant combinations in Arabic can be characterized with reference to the relative dissimilarity of the segments involved. Such effects may well exist in the languages I discuss here, but I have not explored them.

3. I do not attempt to establish the psychological reality of the cooccurrence restrictions under discussion. In offering an analysis that encodes these restrictions in the grammar, I am implicitly assuming that the lexical gaps that provide the evidence for these restrictions have synchronic consequences. For example, I would expect neologisms and loanwords to be adapted to fit existing cooccurrence restrictions. I leave the exploration of this assumption through experimental phonology as a topic for future research.

CHAPTER 2: OUTLINE – DATA AND ANALYSIS

In this chapter, I give the reader an overview of this work, briefly sketching both the data and the analytical approach. The cooccurrence restrictions I discuss fit into four patterns, ranging from strict (Pattern One) to lax (Pattern Four). A simplified picture of these four patterns is given in (1).

(1)

Pattern One Cuzco Quechua, Souletin Basque, Sanskrit

The laryngeal cooccurrence restrictions allow only very dissimilar elements – such as /h/ and ejectives – to cooccur.

Pattern Two Peruvian Aymara, Ofo, Gojri

Only very dissimilar and modestly dissimilar elements can cooccur: for example, /h/ and ejectives, and /h/ and aspirated stops. Identical segments can also cooccur (/h/ with /h/, /t^h/ with /t^h/, /k'/ with /k'/, etc.).

Pattern Three Bolivian Aymara, Old Georgian, Hausa

Only very, modestly, and somewhat dissimilar elements can cooccur: for example, ejectives and /h/, aspirated stops and /h/, heterorganic aspirated stops, and heterorganic ejective and aspirated stops. Identical segments can also cooccur.

Pattern Four Tzutujil, Shuswap

Only extremely similar elements (for example, heterorganic ejectives) are prohibited from cooccurring. Identical segments can also cooccur; Tzutujil is subject to the complete identity effect, while Shuswap is subject to the partial identity effect (explained in 3.4.2).

To the patterns above, we might also add a Pattern Zero, representing languages without categorical laryngeal cooccurrence restrictions.

All of the languages shown in (1) are subject to laryngeal cooccurrence restrictions. In these languages, similar segments (of the set of segments over which laryngeal cooccurrence restrictions operate) do not cooccur, while dissimilar segments do. The point of acceptable similarity is set differently in the different languages.

The similarity scale for relevant segments is presented in (2); this scale is constructed on the basis of the typological statement. Only the dotted and solid lines separate empirically-attested rankings.

(2)

1.	t^h ⇔ ʔ	least similar
2.	t' ⇔ h	
3.	ɗ ⇔ h	
4.	h ⇔ ʔ	
5.	t^h⇔h	somewhat similar
6.	t'⇔k^h	moderately similar
7.	t^h⇔k^h	
8.	ʔ ⇔ ɗ	very similar
9.	ʔ⇔t'	
10.	t' ⇔ ɓ	
11.	ɓ ⇔ ɗ	
12.	t' ⇔k'	
13.	t' ⇔t^h	
14.	h⇔h	maximally similar [=identical]
15.	ʔ⇔ʔ	
16.	t^h⇔t^h	
17.	t'⇔t'	
18.	ɗ ⇔ ɗ	

Here I would like to emphasize that it is similarity among the set of segments over which laryngeal cooccurrence restrictions operate that is at issue. The auditory similarity (or identity) of other segments is not relevant. For example, Cuzco Quechua prohibits two /h/s, identical aspirated stops, and identical ejective stops from cooccurring: *halha, *pʰuspʰu, *t'ult'u. However, forms with identical segments of other types *do* occur: miʎma 'wool', sispa 'near', tatij 'cease, refrain'.

Chapter 3 offers an in-depth presentation of the data from all of the languages listed in (1). For the purposes of this outline, I will only consider data from four languages, each of which is representative of a different pattern: Cuzco Quechua (Pattern One), Peruvian Aymara (Pattern Two), Bolivian Aymara (Pattern Three), and Tzutujil (Pattern Four). The restrictions of these languages were set out in Table 1.2; this chart is reproduced as Table 2.1, below. Recall that a checkmark indicates the absence of a cooccurrence restriction operating over the elements in question; an asterisk indicates the presence of a cooccurrence restriction, and shaded cells indicate that the question is irrelevant for inventory considerations.

	Pattern One: Cuzco Quechua	Pattern Two: Peruvian Aymara	Pattern Three: Bolivian Aymara	Pattern Four: Tzutujil
t^h ⇔ ʔ	✓	✓	✓	
t' ⇔ h	✓	✓	✓	✓
ɗ ⇔ h				✓
h ⇔ ʔ	✓	✓	✓	✓
t^h⇔h	*	✓	✓	
t'⇔k^h	*	*	✓	
t^h⇔k^h	*	*	✓	
ʔ ⇔ ɗ				✓
ʔ⇔t'	*	*	*	✓
t' ⇔ ɓ				✓
t' ⇔k'	*	*	*	*
t' ⇔t^h	*	*	*	
h⇔h	*	✓	✓	✓
ʔ⇔ʔ				✓
t^h⇔t^h	*	✓	✓	
t'⇔t'	*	✓	✓	✓
ɗ ⇔ ɗ				

Table 2.1. Laryngeal cooccurrence restrictions: four patterns.

I provide an Optimality Theory (Prince and Smolensky 1993, McCarthy and Prince 1993, 1994, 1995) analysis of this data, and assume familiarity with the basics of that theory. The full versions of the constraints to which I appeal, and the complete analysis, are presented in chapters 4-6. For the purposes of this overview, I will make use of just four constraints: BEIDENTICAL, *SIMILARITY, IDENTI→O(spread glottis), and IDENTO→I(spread glottis). Definitions of these four constraints are given in (3).

(3)

Constraints (informal)

A. BEIDENTICAL – Segments should be identical. Violations are here counted only between consonant pairs: tata incurs no violations, tanta incurs two, and pata incurs one.

B. *SIMILARITY – Consonants should not be similar. This is actually a family of similarity constraints; each *SIMILARITY constraint sets the degree of similarity that is ruled unacceptable by languages of that pattern, with *SIMILARITY-1 corresponding to languages of Pattern One (Cuzco Quechua, for example), *SIMILARITY-2 corresponding to languages of Pattern Two (Peruvian Aymara, for example), and so on.

A list of the elements prohibited from cooccurring by each of the *SIMILARITY subconstraints is given below. The elements are listed in descending order of similarity; "identity" refers to the cooccurrence of identical elements (i.e., **h** ⇔ **h**, **t^h** ⇔ **t^h**, etc.). Identity resides at the top of the similarity hierarchy because it describes maximal similarity.

Pattern One	**Pattern Two**	**Pattern Three**	**Pattern Four**
Identity	Identity	Identity	Identity
t' ⇔ t^h	t' ⇔ t^h	t' ⇔ t^h	t' ⇔ t^h
t' ⇔ k'	t' ⇔ k'	t' ⇔ k'	t' ⇔ k'
t' ⇔ 6	t' ⇔ 6	t' ⇔ 6	
?⇔t'	?⇔t'	?⇔t'	
? ⇔ 6	? ⇔ 6	? ⇔ 6	
t^h⇔k^h	t^h⇔k^h		
t'⇔k^h	t'⇔k^h		
t^h⇔h			

One *SIMILARITY violation is incurred for every pair of segments violating each *SIMILARITY constraint. For example, hat^{h}a incurs one *SIMILARITY-1 violation; it does not violate *SIMILARITY-2, *SIMILARITY-3, or *SIMILARITY-4. The form ?at'a violates each of the three constraints *SIMILARITY-1, *SIMILARITY-2, and *SIMILARITY-3 once, but does not violate *SIMILARITY-4. A form such as t'at'a – or nana, for that matter – would violate every *SIMILARITY constraint once.

C. IDENTI→O(spread glottis) – If the feature [spread glottis], which characterizes /h/ and aspirated segments, is present on an input segment, then it must also be present on the corresponding output segment (if there is one). The input-output pair /t^haki/-taki would incur one IDENTI→O(spread glottis) violation.

D. IDENTO→I(spread glottis) – If the feature [spread glottis], which characterizes /h/ and aspirated segments, is present on an output segment, then it must also be present on the corresponding input segment (if there is one). The input-output pair /taki/-t^haki would incur one IDENTO→I(spread glottis) violation.

In the analysis presented here, laryngeal cooccurrence restrictions are encoded in the grammar; restrictions on underlying forms are thus rendered unnecessary. The four patterns discussed in this dissertation essentially follow from two conditions in the grammar of the language under consideration: first, whether or not BEIDENTICAL outranks *SIMILARITY, and second, the strength of the active *SIMILARITY constraint. (The active *SIMILARITY constraint is the lowest-ranking *SIMILARITY constraint that outranks some of the laryngeal faithfulness constraints, such as IDENTI→O(spread glottis) and IDENTO→I(spread glottis)). Simplified versions of some constraint rankings that would account for the patterns in question are shown in (4) (a more complete account of these rankings is provided in chapter 5).

(4)

Pattern One	IDENTO→I(sg) >>	*SIMILARITY-1	>>	BEIDENTICAL, IDENTI→O(sg)	
Pattern Two	IDENTO→I(sg) >>	BEIDENTICAL	>>	*SIMILARITY-2 >>	IDENTI→O(sg)
Pattern Three	IDENTO→I(sg) >>	BEIDENTICAL	>>	*SIMILARITY-3 >>	IDENTI→O(sg)
Pattern Four[1]	IDENTO→I(sg) >>	BEIDENTICAL	>>	*SIMILARITY-4 >>	IDENTI→O(sg)

In languages without the identity effect – for example, the Pattern One languages Cuzco Quechua, Souletin Basque, and Sanskrit – the active *SIMILARITY constraint (in these languages, *SIMILARITY-1) outranks BEIDENTICAL. In languages with the identity effect, BEIDENTICAL outranks all of the *SIMILARITY constraints. A few explanatory tableaux are offered below. In these tableaux, an asterisk indicates a constraint violation; exclamation marks signal fatal constraint violations. All cells following fatal constraint violations are shaded, because they are irrelevant for ranking purposes. The pointing finger selects the winning output candidate. Solid column lines indicate that the two constraints separated by that line are ranked as shown; dotted lines indicate that the constraints are unranked.

First, observe that a form that does not violate *SIMILARITY can surface the same in all of the languages under consideration. The tableau in (5) shows how the Pattern One grammar treats an input form containing /ʔ/ and an aspirated stop; both output candidates

[1] The ranking given here is one we might expect of a Pattern Four language with a segment inventory like those found in Cuzco Quechua, Peruvian Aymara, and Bolivian Aymara. No such language is addressed here. In the two Pattern Four languages discussed in this research, the crucially dominated laryngeal faithfulness constraints are IDENTI→O(constricted glottis), IDENTI→O(voice), and IDENTO→I(voice).

involve three BEIDENTICAL violations because both include three pairs of non-identical consonants ({ʔ, m}, {m, p^h/p}, {ʔ, p^h/p}).

(5) Pattern One

	/ʔump^{h}i/	IDENTO→I(sg)	*SIMILARITY-1	BEIDENTICAL	IDENTI→O(sg)
☞	ʔump^{h}i			***	
	ʔumpi			***	*!

The tableau in (6) shows that the same output will surface in a Pattern Two language; Patterns Three and Four work the same way.

(6) Pattern Two

	/ʔump^{h}i/	IDENTO→I(sg)	BEIDENTICAL	*SIMILARITY-2	IDENTI→O(sg)
☞	ʔump^{h}i		***		
	ʔumpi		***		*!

Next, consider a form containing heterorganic aspirated stops, which incurs a *SIMILARITY violation in one language (Peruvian Aymara) but not in another (Bolivian Aymara). Tableau (7) illustrates how the cooccurrence of heterorganic aspirated stops is prevented in Peruvian Aymara.

(7) Pattern Two

	/p^{h}ut^{h}u/	IDENTO→I(sg)	BEIDENTICAL	*SIMILARITY-2	IDENTI→O(sg)
	p^{h}ut^{h}u		*	*!	
☞	p^{h}utu		*		*

A hypothetical input form with heterorganic aspirated stops would sacrifice one laryngeal feature, incurring an IDENTI→O(sg) violation; in doing so it avoids a higher-

ranked *SIMILARITY violation. Other factors, discussed in chapter 6, account for the selection, in this language, of pʰutu over the alternative candidate *putʰu. This tableau shows how, in Optimality Theory, cooccurrence restrictions can be encoded in the grammar, rather than in a set of conditions levied on underlying forms (for example: "underlying forms may not contain similar segments of a particular type", or "laryngeal features must be floating in underlying representations"); see Smolensky 1996 on the advantages of allowing underlying forms to be rich.

Unlike Peruvian Aymara, Bolivian Aymara allows the cooccurrence of heterorganic aspirated stops. This follows from *SIMILARITY being a weaker constraint (i.e., a constraint that penalizes fewer things for cooccurring) in Bolivian Aymara than it is in Peruvian Aymara. The effect of the different parameterizations of *SIMILARITY can be seen by comparing the tableau in (7) to that in (8).

(8) Pattern Three

	/pʰutʰu/	IDENTO→I(sg)	BEIDENTICAL	*SIMILARITY-3	IDENTI→O(sg)
☞	pʰutʰu		*		
	pʰutu		*		*!

The weaker version of *SIMILARITY implemented in Bolivian Aymara allows heterorganic aspirated stops to cooccur.

Finally, consider the case of identical aspirated stops, which are ruled out in Pattern One languages, but are permitted in Pattern Two languages (among other languages). Tableau (9) illustrates the case of a Pattern One language.

(9) Pattern One

	/pʰapʰi/	IDENTO→I(sg)	*SIMILARITY-1	BEIDENTICAL	IDENTI→O(sg)
☞	pʰapi			*	*
	pʰapʰi		*!		

The tableau above shows that a form with identical aspirated stops will incur a fatal *SIMILARITY-1 violation in a Pattern One language. The winning candidate in (9) is pʰapi – a form that has sacrificed one aspiration feature in order to avoid a *SIMILARITY violation. In Pattern Two languages, by contrast, the candidate that preserves both of its [spread glottis] features is preferred: (10) shows that pʰapʰi satisfies high-ranked BEIDENTICAL, and unlike *papi it incurs no IDENTI→O(sg) violations.

(10) Pattern Two

	/pʰapʰi/	IDENTO→I(sg)	BEIDENTICAL	*SIMILARITY-2	IDENTI→O(sg)
☞	pʰapʰi			*	
	pʰapi		*!		*
	papi			*	**!

Regarding tableaux (9) and (10), a parallel argument could be made with respect to forms with identical ejective stops. This would account for the presence of forms such as k'ink'u 'clay' and t'ult'u 'stubble field' in Peruvian Aymara, and for the absence of such forms in Cuzco Quechua. Because BEIDENTICAL is formulated as a single constraint (rather than as separate constraints of the form "aspirated segments must be identical", "ejective segments must be identical", etc.), identical segments of the set over which laryngeal cooccurrence restrictions operate will typically pattern in the same way within any language. This is how the monolithic nature of the identity effect is modeled.

I omit the tableaux for identical ejective stops here, as I omit discussion of forms with glottal stops, implosives, or /h/ in combination with another aspirated segment. A complete analysis of these phenomena can be found in chapters 4 and 5. At this point, the reader should have a general idea of the data patterns under study, and of the analytical approach taken here.

CHAPTER 3: DATA

The data for this chapter were gathered by searching through dictionaries by hand. In some cases, I could not be sure of the morphological structure of the entry under consideration; this is a real concern since most of the restrictions discussed here operate at the morpheme level. My solution was to err on the side of generosity in generating lists of apparent counterexamples. In any case, I may have missed some forms; the numbers of apparent counterexamples cited here should be viewed as approximately correct.

3.1 Pattern One

Pattern One cooccurrence restrictions are the most restrictive; they prohibit even rather dissimilar segments from cooccurring. Table 3.1 sets out which segments can cooccur within morphemes (or, in the case of Sanskrit, within roots) and which cannot. Recall from the first chapter that **t'** represents an ejective segment at any place of articulation, **t'⇔t'** designates the cooccurrence of identical ejectives, **t'⇔k^h** describes the cooccurrence of an ejective with a heterorganic aspirated stop, and so on.

The third column in the table ("Attested in . . . ") indicates which languages have inventories that include both of the segments under consideration. These are the languages that have the possibility of manifesting the cooccurrence restriction in question. The abbreviations for the different Pattern One languages are as follows: "CQ" = Cuzco Quechua, "SB" = Souletin Basque, and "Skt" = Sanskrit. Note that some cells in the third

column are blank: these are cells that are irrelevant for inventory reasons in *all* of the languages under consideration. For example, none of the languages discussed in this section contain implosives, so the asterisk in the second column of the ɓ⇔ɗ row actually represents an extrapolation from existing data. Such rows are included in Table 3.1 to facilitate the comparison of restrictions in languages of different patterns.

	Pattern One	Attested in ...
t^h⇔ʔ	✓	CQ
t'⇔ h	✓	CQ
t^h⇔ɓ	✓	
ɗ⇔h	✓	
h⇔ʔ	✓	CQ
t^h ⇔h	*	CQ, SB, Skt
$d^ɦ$⇔h	*	Skt
t'⇔k^h	*	CQ
t^h ⇔ $g^ɦ$	*	Skt
$d^ɦ$⇔$g^ɦ$	*	Skt
t^h ⇔k^h	*	CQ, SB, Skt
t^h ⇔ $d^ɦ$	*	Skt
ʔ ⇔ɗ	*	
ʔ⇔t'	*	CQ
t'⇔ɓ	*	
ɓ⇔ɗ	*	
t'⇔k'	*	CQ
t'⇔ t^h	*	CQ
h⇔h	*	CQ, SB, Skt
ʔ⇔ʔ	*	
t^h⇔t^h	*	CQ, SB, Skt
$d^ɦ$⇔$d^ɦ$	*	Skt
t'⇔t'	*	CQ
ɗ⇔ɗ	*	

Table 3.1. Laryngeal cooccurrence restrictions of Pattern One languages.

In addition to the data shown in Table 3.1, all Pattern One languages allow the cooccurrence of voiceless unaspirated stops with segments that are identical except in their laryngeal features: t^{h}ata, k'ika, etc. For reasons that will be made clear in chapters 4 and 5, this characteristic seems to accompany the absence of the complete identity effect (i.e., the lack of forms containing identical aspirated stops, identical ejectives, etc.). It is the sole respect in which the Pattern One restrictions are weaker than those of the Pattern Two, Three, and Four languages.

3.1.1 Cuzco Quechua

3.1.1.1 Cuzco Quechua - language basics

Various dialects of Quechua are spoken by four to five million people in Colombia, Ecuador, Peru, Bolivia, Chile, and Argentina (Crapo and Aitken 1986). The dialect of Quechua discussed in this study is Cuzco Quechua (Quechuan:Quechua A:Southern; Grimes 1988), spoken by about 1,250,000 people in the Cuzco and Arequipa regions of Peru (Grimes 1988, citing Cusihuaman 1976a). The segment inventory of Cuzco Quechua is given in (1).

(1)

p	t	tʃ	k	q	
p'	t'	tʃ'	k'	q'	ʔ
p^h	t^h	$tʃ^h$	k^h	q^h	
	s	ʃ			h
	ɾ				
m	n	ɲ			
	l	ʎ			
w		j			

Glottal stop appears only in word-initial position; there are no vowel-initial words in the language (Rowe 1950, citing J.P. Harrington as the source of this observation).[1]

The vowel system of Cuzco Quechua is /i/e/a/o/u/; /e/ and /o/ occur mostly in Spanish loanwords and as a result of vowel lowering near uvulars. Syllables are of form (C)V(C), and most roots are disyllabic. The language is suffixing; stress is penultimate.

[1] Carenko 1975 suggests that glottal stop is phonetically optional. However, the basis of this statement seems to be that although some researchers have commented on the presence of glottal stop, others have not mentioned it. Carenko does not cite any actual observations of glottal stop-initial forms in alternation with vowel-initial forms.

3.1.1.2 Cuzco Quechua - cooccurrence restrictions

Various scholars have noted the existence of cooccurrence restrictions in Cuzco Quechua (Rowe 1950, Torero Fernandez de Cordova 1964 [not seen], Orr and Longacre 1968, Carenko 1975). A complete list of the restrictions is given in (2); all data are taken from Hornberger and Hornberger 1983, a dictionary of about 2000 forms. Like other Quechua dictionaries, Hornberger and Hornberger does not transcribe glottal stop; in accord with the statement in 3.1.1.1, I have added it to all forms that are listed in the dictionary as being vowel-initial.

(2)

a. There is only one ejective per morpheme: t'anta 'bread'; *t'ant'a, *t'ank'a.
b. There is only one aspirated stop per morpheme: q^hotʃu 'flock', *q^hop^{h}u, *q^hoq^hu.
c. There is only one /h/ per morpheme: *halha.
d. Aspirated and ejective stops may not cooccur: *t'ap^{h}a, *t'at^{h}a.
e. Aspirated stops and /h/ may not cooccur: *hap^{h}a.
f. Glottal stop may not cooccur with ejectives: *ʔat'a.

Although /h/ does not cooccur with aspirated stops, it does cooccur with ejectives (hump'i 'sweat' and many others; the vast majority of Cuzco Quechua /h/s are found in initial position, but we do find two examples of morphemes with ejective segments followed by /h/: tʃ'oho 'cough', k'ahaj 'to split'), and of course with voiceless unaspirated stops: huku 'owl', hanku 'raw'. Voiceless unaspirated stops may also cooccur: kiku 'a yellow flower', qarqoj 'to eject, dismiss', takij 'to sing'.

Cuzco Quechua allows morphemes with homorganic aspirated and plain stops, and morphemes with homorganic ejective and plain stops: t^hanta 'old, used, worn out',

q'oqa 'coarse, brute'. The Aymara dialects (as will be seen in 3.2.1 and 3.3.1) are examples of languages that disfavor the occurrence of such forms; this is the sole respect in which the Pattern Two and Three restrictions are not a subset of the Pattern One restrictions.

Ejective and aspirated stops appear only in Cuzco Quechua root morphemes; they do not appear in suffixes. In any case, the treatment of compounds (discussed briefly in the following section) suggests that the domain of laryngeal cooccurrence restrictions is the morpheme; no word-level alternations are observable.

Finally, if an ejective or aspirated stop appears in a morpheme, it will be the first, or "leftmost" stop in the morpheme: q^{h}otʃu, *qotʃ^{h}u. I refer to this as the "leftness effect". This effect is occasionally formulated (in the works cited above) as applying to the leftmost *onset* stop. However, Rowe 1950 observes that the segments written as coda stops in Cuzco Quechua are in fact fricatives: orthographic *aknu* 'red-violet color', for example, is phonetically [ʔaxnu]. Coda "stops" (which I will continue to transcribe as stops, following the practice of all researchers working on the language) are therefore ineligible for aspiration and glottalization, and the leftness effect can be formulated in terms of the leftmost *stop* in the morpheme.

3.1.1.3 Cuzco Quechua - descriptive details

Hornberger and Hornberger include three exceptions to the leftness effect: tonqʰo 'poorly done work', tʃust'ij 'to skin', and puskʰaj 'to spin (into thread)'. However, two of these morphemes have regular alternate forms: 'to skin' is also found in Hornberger and Hornberger as ʎust'ij, and 'to spin' is found as puskaj in Cusihuaman 1976b, a dictionary of the Cuzco-Collao dialect region.

As with the leftness effect, the cooccurrence restrictions given in (2) are not exceptionless. There are two forms in Hornberger and Hornberger that contain two /h/s, in violation of (2c) above. However, one and perhaps both of these forms (hahatij 'to pant', hahatʃu/hahaʎu 'male duck') are onomatopoeic.

A close study of Hornberger and Hornberger also reveals two forms that violate (2f): ʔusut'a 'sandals' and ʔajtʃ'awij 'embalm'. The second of these two forms could well be bimorphemic (cf. ʔajtʃa 'meat, flesh'), in which case it would not constitute a violation of (2f).

In the rest of this work, I will disregard this small handful of exceptions to the observations made in 3.1.1.2.

Hornberger and Hornberger also reveal two sets of spurious counterexamples. First, two forms found in that dictionary have more than one ejective: tʃ'atʃak'umu 'large, curly-trunked, pink-flowered South American tree' and tʃ'iʎik'utu 'grasshopper'. Both, however, appear to be compounds composed of unexceptional roots (cf. tʃ'atʃa

'curly', k'umu 'bent over', and k'utuj 'to bite with incisor teeth'; I couldn't find a root corresponding to tʃ'iʎi).

Second, several monomorphemic forms also appear to have more than one aspirated or ejective stop. Examples are provided in (3).

(3)

k'ap^{h}ka	half-cooked
ʎip^{h}t'a	substance placed in mouth when chewing coca leaves
lap^{h}t'aj	to lap water like a dog
ʎip^{h}tʃ'ij	snip off with the finger nails

Orr and Longacre 1968 suggest that forms such as these show that Cuzco Quechua allows more than one aspirated or ejective stop per word. However, a search of Hornberger and Hornberger reveals that one of the stops in question is always coda *p^h*. Two pieces of evidence suggest that this transcription by Hornberger and Hornberger actually designates [ɸ], the coda allophone corresponding to onset [p].

First, Hornberger and Hornberger transcribe the following set of sounds in coda position: {w, j, n, m, l, ʎ, ɾ, s, q, k, t, p^h}. Assuming that the sound transcribed *p^h* is some variant of a voiceless bilabial, and that Hornberger and Hornberger are transcribing coda fricatives as stops, we would expect to find /p/ rather than /p^h/ here. In that case, the set of allowable coda consonants would include all sonorants except for /ɲ/ and /h/, and fricated versions of all plain stops. In other words, some allophone of /p/ is the expected segment on phonotactic grounds.

Second, observations made by other scholars who have worked on the language also suggest that these segments cannot be aspirated bilabial stops. As mentioned above, Rowe 1950 states that in Cuzco Quechua, the syllable-final allophones of stops are fricatives: for example, [ɸ] is found in place of [p]. Mannheim 1991 states that syllable-final plain stops in the Ayacucho-Chanka dialect correspond to fricatives in the Cuzco-Collao dialect region, and offers the cognate pair <u>ʎipta</u> ~ <u>ʎiɸt'a</u> 'ash lime chewed with coca' in order to illustrate this relation. The comparison of <u>ʎiɸt'a</u> with *ʎip^{h}t'a* (cited in (3)) suggests that Hornberger and Hornberger transcribed [ɸ] as *p^h*. I conclude that Hornberger and Hornberger transcribe coda allophones of stops – which have incomplete oral closure – using the symbols *p^h*, *t*, *k*, and *q*; *p^h* represents [ɸ] rather than [p^h]. [2]

I conclude that Cuzco Quechua observes the restrictions described in 3.1.1.2.

[2] I should note here that the Hornbergers' dictionary includes two forms that do not coincide with my understanding of their transcription system. These forms are *hapq'ey* 'to disinter' and *mit^{h}may* 'to exile', for which I would expect *hap^{h}q'ey* (<u>hapq'ej</u>) and *mitmay* (<u>mitmaj</u>). I will disregard this minor inconsistency.

3.1.2 Souletin Basque

3.1.2.1 Souletin Basque - language basics

Basque, a language isolate, is spoken by over 700,000 people in Spain and France (Grimes 1988, relying on census data from the 1970's). Souletin Basque is spoken on the northern side of the Pyrenees; I was unable to find a modern estimate of the number of speakers of this dialect. The consonant inventory of Souletin Basque is given in (4) (Hualde 1993).

(4)

p	t	ʈs	ts	tʃ	c	k		
pʰ	tʰ					kʰ		
b	d	(ɖʐ)			ɟ	g		
f	s	ʂ		ʃ			h	ɦ
	z	ʐ		ʒ				
m	n				ɲ			
	l				ʎ			
	r							

The voiced retroflex fricative /ʐ/ and the voiced apico-alveolar fricative /z/ appear mostly in loanwords. The voiced retroflex affricate /ɖʐ/ occurs only intervocalically in loanwords.

The vowels of Souletin Basque are /a/e/i/o/u/y/; vowels may be long or short, and may also surface as nasalized.[3] Stress is contrastive and falls on the final or (more commonly) the penultimate syllable.

3.1.2.2 Souletin Basque - cooccurrence restrictions

The cooccurrence restrictions of Souletin Basque were first noted by Lafon 1958; they are given in (5).[4]

(5)

a. There is only one aspirated stop per morpheme: *k^{h}at^{h}o, *k^{h}ak^{h}o.
b. There is only one laryngeal fricative per morpheme: *haha, *hiha.
c. Aspirated stops do not cooccur with laryngeal fricatives: *hat^{h}a, t^{h}aha.

Again, as in Cuzco Quechua, identical voiceless unaspirated segments may cooccur (kyky'so 'flea', po'pyly 'nation'), as may homorganic voiceless aspirated and

[3] Hualde proposes that nasalized vowels are derived. This analysis is supported by the observation that nasal vowels are typically adjacent either to nasal consonants or to other nasalized vowels; oral vowels do not occur in these positions. In other words, if [nasal] spread bidirectionally but spreading was blocked by oral consonants, then almost all occurrences of nasal vowels would be accounted for. However, the existence of a few forms such as goñi'ũ 'reddish mushroom' and ar'dũ 'wine' suggest that distinctively nasalized vowels may in fact exist (Hualde accounts for these forms with a rule of final nasal consonant deletion that has lexically-marked exceptions). In any case, this point is not important for the data discussed here; I follow Hualde's practice of leaving nasalized vowels unmarked except in the rare cases where their occurrence is not predictable from context.

[4] Interestingly, Michelena 1976:211 points out that the restriction against two occurrences of aspiration within a word does not hold true of medieval Basque, at least as that language is written. This point could be pursued in order to discover more about the diachronic course of laryngeal cooccurrence restrictions.

voiceless unaspirated segments ('k^{h}alka 'to stuff', 'k^{h}ako 'to hook'). Souletin Basque adds a new wrinkle, however, in that voiced obstruents are also present in its inventory; identical voiced stops are permitted to cooccur ('dyynda 'thunder', 'gogor 'hard'), as are homorganic voiceless aspirated and voiced segments ('gark^{h}otʃ 'nape', k^{h}o'gan 'beehive'), and voiceless unaspirated and voiced segments (de'hota 'sterile', 'gaʃki 'bad').

Michelena 1976 reports a leftness effect operating upon aspiration; this leftness effect is similar to that seen in Cuzco Quechua (described in 3.1.1) and the Aymara dialects (described in 3.2.1 and 3.3.1). In forms that include both voiceless aspirated and voiceless unaspirated stops, the aspirated stop must precede the unaspirated stop (due to the restrictions in (5), of course, aspiration cannot occur on both stops). Thus, words of form 'k^{h}alka 'to stuff' are found, but words of form kalk^{h}a are not. As Hualde points out, however, the segment */h/ can* occur after a voiceless unaspirated stop: tra'hel 'person with crippled legs', ke'hela 'screen, lattice-work door'.[5]

---

[5] The morpheme structure constraints in (4) also apply at the word level: 'hil 'dead' + 'heri 'country, land' > 'ilheri 'graveyard' (*hilheri). Michelena 1976 instructs us that in such cases, it is usually the first of the two aspirates that deaspirates. It is interesting that although deaspiration usually targets the first of two aspirates (as in Grassmann's Law in Greek and Sanskrit, deaspiration – and also deglottalization – in the Salishan languages, deaspiration in Ofo, etc.), when the location of aspiration within a morpheme is restricted, we always find a *leftness* effect (as in Cuzco Quechua, Peruvian Aymara, and Bolivian Aymara), rather than a rightness effect. It would be instructive to look at Basque compounds of shape C^h ... C^h, to see whether it is the first or second aspirate that deaspirates. Unfortunately, Michelena does not supply the relevant forms.

Not all words with two aspirated segments undergo deaspiration. Compare 'ilheri with hilhotʂ (hil 'dead' + hotʂ 'cold') 'rigor mortis'. It may be the case that the internal structure of polymorphemic words that violate the laryngeal cooccurrence restrictions is more strongly felt. A Basque listener would know from the antepenultimate stress on

3.1.2.3 Souletin Basque - descriptive details

I checked the cooccurrence restrictions observed by Lafon against Larrasquet 1939, a lexicon that contains about 6500 entries broken down by morphemes (not all forms are labeled as belonging to the Souletin dialect). There are no monomorphemic forms that violate the restrictions in (5).

I also searched for counterexamples to the leftness effect mentioned in the preceding section. There are none listed in the dictionary: in all entries involving both a voiceless aspirated and a voiceless unaspirated stop, the aspirated stop precedes the unaspirated stop. (Michelena 1976 writes that this holds true only of words in which the relevant stops occur in the first two syllables. However, the Larrasquet dictionary contains no counterexamples to the broader statement made above.)

3.1.3 Sanskrit

3.1.3.1 Sanskrit - language basics

My main sources for Sanskrit (Indo-European:Indo-Iranian:Indo-Aryan) data are Whitney 1941, 1945.

The segment inventory of the language is shown in (6).

'ilheri that the word is a compound (I thank Gorka Elordieta for pointing this out to me), but hilhots might still feel more wordlike.

(6)

p	t	ʈ	c	k	
p^h	t^h	$ʈ^h$	c^h	k^h	
b	d	ɖ	ɟ	g	
$b^ɦ$	$d^ɦ$	$ɖ^ɦ$	$ɟ^ɦ$	$g^ɦ$	
	s	ʂ	ç		h
m	n	ɳ	ɲ	ŋ	
v	l		j		
	ɻ				

Sanskrit syllable nuclei are transcribed as follows: /a/i/u/a:/i:/u:/e/o/a:i/a:u/ɻ/ɻ:/l̩/.

In the data presented in these sections, I transcribe what is by Whitney's account the retroflex approximant /ɻ/ as 'r', and I transcribe the palatal glide /j/ as 'y', in accord with the transcription systems used by earlier researchers.

3.1.3.2 Sanskrit - cooccurrence restrictions

Sanskrit is one of the languages for which Grassmann's Law was formulated.[6] The initial observation was stated in diachronic terms: the first of two aspirates beginning successive syllables deaspirates, as does the initial aspirate in a syllable that both opens and closes with aspirates (Lehmann 1966:1, citing Grassmann 1863).

The Sanskrit restrictions are restated in (7).[7]

---

[6] The other Grassmann's Law language was of course Attic, or ancient Greek. This language appears to have been a Pattern Two language without the identity effect. Due to time constraints, I was unable to include that language in this study.

[7] A similar set of restrictions is found in another Indo-Aryan language, the Rajasthani language Harauti. The Harauti segment inventory is similar to that of Sanskrit; one relevant respect in which the languages differ is that Harauti possesses the voiceless sonorants /m̥/n̥/l̥/. Allen 1957 describes the cooccurrence restrictions of this language.

(7)

a. There is only one aspirated stop per root, whether the stops in question are voiceless or voiced.
b. There is only one /h/ per root.
c. Aspirated stops, whether voiced or voiceless, do not cooccur with /h/.

These restrictions operate at the root level, and at the level of root plus reduplicated material. For example, the perfect forms of the roots p^{h}al 'burst' and d^{ɦ}auk 'approach' are pap^{h}ala and dud^{ɦ}auka, respectively (*p^{h}ap^{h}ala, *d^{ɦ}ud^{ɦ}auka; Anderson 1970). The laryngeal cooccurrence restrictions do not operate over larger domains: aspirated elements can coexist in roots and suffixes. Consider, for example, bib^{ɦ}rt^{h}a, which is the 2nd pl. active present of b^{ɦ}r 'bear' (*bibrt^{h}a), and dad^{ɦ}a:t^{h}e, which is the 2nd dual middle present of d^{ɦ}a: 'put' (*dada:t^{h}e; Anderson 1970).

Sanskrit does provide examples of combinations of other types of segments. For example, voiceless aspirated segments can cooccur with voiceless unaspirated and voiced segments: katt^h 'boast', k^{h}a:d 'chew' (I was unable to find examples where the relevant

---

That author states that none of the aspirated segments can cooccur within morphemes, where "aspirated segments" includes /h/, voiceless aspirated stops, voiced aspirated stops, and voiceless sonorants. Thus, the Harauti restrictions are just like those of Sanskrit, except that voiceless sonorants exist and participate in the cooccurrence restrictions. In addition to these restrictions, some ordering restrictions apply, just as in Cuzco Quechua and the Aymara dialects. Voiced aspirated obstruents, /h/, and aspirated (i.e. voiceless) sonorants appear only in word-initial position. Voiceless obstruents cannot precede aspirated segments in a word. In other words, aspiration will occur either in initial position, or on the first voiceless element of the word.

Unfortunately, Allen gives virtually no actual data from Harauti. I will not discuss this language further here.

segments are homorganic – /tt^h/ probably designates a geminate – but aspirated segments are rather uncommon). We also find roots containing voiceless unaspirated segments cooccurring with other voiceless unaspirated segments, and with voiced segments (kamp 'tremble', ka:ŋks 'desire', pad 'go', trd 'split, bore'), roots with two voiced stops (garj 'roar', da:/dad 'give'), and roots containing both voiced aspirated segments and voiceless unaspirated segments or voiced segments (g^{ɦ}at 'strive', grd^{ɦ}, 'be greedy', d^{ɦ}a:/dad^{ɦ} 'put'; again, I found no examples of the cooccurrence of voiced aspirated stops with homorganic voiced stops, but this may be due to the infrequency of aspirated segments).

3.1.3.3 Sanskrit - descriptive details

Whitney 1945 includes only two roots with more than one aspirated element: hu:rc^h 'fall away', and heʈ^h (gloss unspecified). The first is noted as related to the unexceptional root hvr/hru/hur 'be or make crooked', but does appear to constitute a counterexample to the restrictions listed above. Whitney (p. 208) writes that heʈ^h probably arises from a misreading of -ḍaka.

Sanskrit includes a set of alternations that has served as the subject of many phonological analyses (for example, Anderson 1970, Phelps and Brame 1973, Sag 1974, 1976, Hoard 1975, Phelps 1975, Schindler 1976, Borowsky and Mester 1983, Lombardi 1991, and Broe 1993). The alternations concern a handful of roots in which aspiration shifts from one consonant in the root to another. The migration of aspiration within the root is determined by the environment in which the root appears (compare, for example,

bod^{ɦ}ati (3rd sg. present indicative of 'to know') and b^{ɦ}otsyati (3rd sg. future)). Although a complete account of laryngeal cooccurrence restrictions should address the Sanskrit alternations, I do not discuss these facts here.

3.1.4 Pattern One – summary

The three languages discussed in the preceding sections – Cuzco Quechua, Souletin Basque, and Sanskrit – all bear out the Pattern One restrictions. Aspirated segments do not cooccur within roots, nor do ejectives; glottal stop does not cooccur with itself or with ejectives in any of these languages. There are also no cases of ejective segments cooccurring with aspirated segments.

None of these languages have inventories that include implosives, but if such languages were found, we would expect them to include restrictions against the cooccurrence of implosives with ejectives, glottal stop, and other implosives.

The identity effect does not exist in Pattern One languages. This means that it is not the case (as it is in most of the languages discussed in the following sections) that similar segments of the set investigated here – /h/, /ʔ/, voiceless aspirated, voiced aspirated, ejective, or implosive segments – can cooccur just in case they are identical.

Finally, all of these languages allow the cooccurrence of voiceless aspirated segments with segments that differ in laryngeal setting but are otherwise identical: for example, k'ika, k^{h}ika, etc. This is significant since it is not generally true of languages with the identity effect (i.e., languages in Patterns Two through Four).

3.2 Pattern Two

Pattern Two cooccurrence restrictions are less strict than Pattern One restrictions; they can be informally characterized with the statement that somewhat similar elements are prohibited from cooccurring. The following sections cover three languages that fit the Pattern Two cooccurrence restrictions. A schematic showing which segments can and which segments cannot cooccur within morphemes in these languages is provided in Table 3.2. With respect to the third column of this table, the abbreviations are as follows: "PA" represents Peruvian Aymara, "Ofo" indicates Ofo, and "Goj" stands for Gojri.

Notice that Table 3.2 suggests that Ofo may have lacked forms of type $\mathbf{t^h \Leftrightarrow t^h}$. Given the available data, we cannot be certain whether or not such forms existed in this language (see 3.2.2.4 for more information). In the analysis presented in chapters 4 and 5, I argue that these forms must have been absent in the language.

	Pattern One	Attested in ...
t^h⇔ʔ	✓	PA
t'⇔ h	✓	PA
t^h⇔ɓ	✓	
ɗ⇔h	✓	
h⇔ʔ	✓	PA
t^h⇔h	✓	PA, Ofo, Goj
$d^{ɦ}$⇔h	✓	Goj
t'⇔k^h	*	PA
t^h ⇔ $g^{ɦ}$	*	Goj
t^h⇔k^h	*	PA, Ofo, Goj
$d^{ɦ}$⇔ $g^{ɦ}$	*	Goj
t^h ⇔ $d^{ɦ}$	*	Goj
ʔ ⇔ɗ	*	
ʔ⇔t'	*	PA
t'⇔ɓ	*	
ɓ⇔ɗ	*	
t'⇔k'	*	PA
t'⇔t^h	*	PA

h⇔h	✓	PA, Ofo
ʔ⇔ʔ	✓	
t^h⇔t^h	✓ / *	PA, Goj / Ofo
$d^{ɦ}$⇔$d^{ɦ}$	✓	Goj
t'⇔t'	✓	PA
ɗ⇔ɗ	✓	

Table 3.2. Laryngeal cooccurrence restrictions of Pattern Two languages.

3.2.1 Peruvian Aymara

3.2.1.1 Peruvian Aymara - language basics

Aymara (Aymaran; Grimes 1988[8]) is spoken by about three million people, mainly in Peru and Bolivia (Hardman 1985, citing an estimate from El Instituto de Lengua y Cultura Aymara). My data sources on the Aymara dialect described in this section are Deza Galindo 1989 and Ayala Loayza 1988; my source for the Aymara dialect described in 3.2.2 is De Lucca 1987.

Briggs 1985 defines morphophonemic, morphological, and lexical characteristics of various Aymara dialects. Based on the lexical tests I was able to apply, my sources describe dialects that fit her 'Northern' grouping, composed of dialects spoken near Lake Titicaca. However, the Deza Galindo 1989 and Ayala Loayza 1988 dictionaries (both published in Peru), although not identical for all forms, describe dialects that vary in regular ways from those included in the De Lucca 1987 dictionary (published in Bolivia). Although I refer to the data in Deza Galindo and Ayala Loayza as "Peruvian Aymara", and to the data in De Lucca as "Bolivian Aymara", these designations are used for the

[8] There is a debate among Andeanists over whether or not Quechua and Aymara are descended from a common ancestor language. See Orr and Longacre 1968 for advocacy of the common origin hypothesis; Hardman 1985, Briggs 1985, Mannheim 1991, and Landerman 1994 all include discussions of the debate. Whether or not these languages are related, the existence of cooccurrence restrictions of the Quechua and Aymara types in other languages – for example, Souletin Basque, Gojri and Hausa – indicates that neither genetic relatedness nor the geographical proximity of these languages is essential to the typology established here.

sake of simplicity. I am not claiming that the dialects I discuss here respect national boundaries.

The consonant inventory of Peruvian Aymara is shown in (8).

(8)

p	t	tʃ	k	q	
p'	t'	tʃ'	k'	q'	
p^h	t^h	$tʃ^h$	k^h	q^h	
	s	ʃ[9]			h
	ɾ				
m	n	ɲ			
	l	ʎ			
w		j			

As in Cuzco Quechua, the vowel system of Peruvian Aymara is /i/u/a/e/o/; /e/ and /o/ occur in Spanish loanwords and as a result of vowel lowering near uvulars. Most roots are of form CVCV(CV); CVCCV and VC(C)V roots are also common, and suffixes are typically of CV shape (Martin 1988). Suffixes in this language, unlike in Cuzco Quechua, may contain aspirated and ejective segments. The language is suffixing; stress is penultimate.

[9] This sound is present in only one of the two dictionaries I consulted (Deza Galindo 1989). All words containing this segment are also found with alternate pronunciations in which /s/ replaces /ʃ/.

3.2.1.2 Peruvian Aymara - cooccurrence restrictions

Cooccurrence restrictions in the Aymara dialects are discussed by Hardman et al. 1974, Adelaar 1986, and Landerman 1994. The differences between the dialects I have labeled "Peruvian" and "Bolivian" (compare (9) with (21)) were first described in MacEachern 1997.

A list of the cooccurrence restrictions of Peruvian Aymara is given in (9). All data is taken from Ayala Loayza 1988 and Deza Galindo 1989; both are dictionaries of around 9000 forms.

(9)

a. There is only one ejective per morpheme, unless the ejectives are identical: k'ink'u 'clay,' t'ult'u 'stubble field of barley or quinoa'; *t'ank'a.
b. There is only one aspirated stop per morpheme, unless the aspirated stops are identical: $p^h usp^h u$ 'boiled beans', $tʃ^h atʃ^h u$ 'tattered'; *$p^h atʃ^h u$.
c. Aspirated and ejective stops do not cooccur: *$t'ap^h a$, *$t'at^h a$.
d. Vowel-initial forms do not include ejectives: *at'a.
e. Morphemes with homorganic aspirated and plain stops, and morphemes with homorganic ejective and plain stops, are rare: *$t^h ata$, *t'ata.

In Peruvian Aymara, there are no restrictions on what may cooccur with /h/. As in Cuzco Quechua, the laryngeal fricative may cooccur with ejectives (heʎq'e 'eye secretions'). However, unlike Cuzco Quechua, Peruvian Aymara also allows the cooccurrence of two /h/s and of /h/ and aspirated stop: halha 'ominous', $hat^h i$ 'heavy', $p^h ahi$ 'moderate, lukewarm'.

With respect to (9c), I must note that I was unable to find a statement in the literature to the effect that apparently vowel-initial Aymara forms actually begin with

glottal stop. It is because of this that I omitted /ʔ/ from the segment inventory given in (8); I will also refrain from adding it to the transcriptions given here, but the approach I take in my analysis is that these forms are in fact glottal stop-initial.

Like Cuzco Quechua, Peruvian Aymara is subject to the leftness effect: aspirated and ejective features appear on the leftmost available stop in the morpheme. As in Cuzco Quechua, aspirated and ejective segments cannot appear in coda position within roots (Martin 1988; clusters that are not permitted within roots do occur as a result of affixation and morphophonological processes).

3.2.1.3 Peruvian Aymara - descriptive details

There are some exceptions to the cooccurrence restrictions recorded above. In the lists that follow, I consider only forms that have three or fewer syllables; forms of four or more syllables are overwhelmingly likely to be polymorphemic.

There is one form listed in Ayala Loayza and Deza Galindo that might be monomorphemic and that includes heterorganic ejective stops, in violation of restriction (9a): q'iwtʃ'ukta 'make signs at a distance'. However, the structure of this form (three syllables, with a CC cluster separating the first and second syllables) suggests that it is bimorphemic, although I was not able to decompose it.

With respect to (9b), which states that heterorganic aspirated segments do not cooccur, we find just one counterexample: qʰatʃʰua 'dance'.

Restriction (9c) states that aspirated and ejective segments do not cooccur. We find eight apparent exceptions. The exceptions are q^{h}oʎq'u 'chubby', q'uq^{h}ata 'in front, forehead', q^{h}oq'a 'mold, rust', p^{h}ultʃ'i 'to wrap, muffle', p^{h}uʎtʃ'u 'wide', tʃ'ap^{h}ɾa 'bough, branch', q'ap^{h}i 'fragrance', and q^{h}aɾajp'u 'tomorrow night' (this last form could well be bimorphemic).

Restriction (9d), as I interpret it, prohibits glottal stop from cooccurring with ejectives. We find 12 examples of apparently monomorphemic Peruvian Aymara forms that contain ejectives and are recorded in the dictionaries with initial vowels (recall that I am suggesting that these forms are in fact glottal stop-initial). Of this list of counterexamples, one is a loanword (asut'i 'whip'; cf. Spanish *azote*) and three have regular alternate forms (isk'a/hisk'a 'little', iʎap'i/iʎapa 'firing, shooting', and amawt'a/amawta 'master, thinker'). The eight non-alternating, apparently native forms are shown in (10).

(10)

usut'a	sandal
atʃt'ajma	to contain, support
apt'i	loss of money
aq'e	cave
aɾtʃ'ukiɲa	to alarm, shout, insult
aɾq'e	struggle
iʎap'a	thunderbolt
uɾq'u	black woolen tunic

In the two Peruvian Aymara dictionaries, we find around two dozen potentially monomorphemic forms that include homorganic ejective and voiceless unaspirated stops, or homorganic aspirated and voiceless unaspirated stops, and that do not have regular alternants (i.e., alternate forms that do not violate (9e)). A complete list of these forms is given in (11).[10]

[10] In a few instances, the orthographies used by Deza Galindo and Ayala Loayza are opaque to me. I have transcribed both Deza Galindo's *jh* sequences and Ayala Loayza's *j'* sequences as simply 'h'; both correspond to simple *j* (/h/) symbols in the other dictionary. The forms in (11) that involve these transcription decisions are the words for 'tattered' (orthographic *chhojhchi*), 'fried food made with quinoa flour' (*thajhti, thojhto*), and 'eaves' (*chhoj'chu*); the first forms are from Deza Galindo, while the last is from Ayala Loayza.

(11)

tʃʰitʃiki	in tumult
tʃʰohtʃi	tattered
tʃʰuʎtʃu	many
k'aʎka	crack
k'aʎku	sour
q'aʎqa	flirtatious, malicious
q'enqo	labyrinthine, zigzag
qʰanqa	nasal, twangy
antʃʰitʃa/antʃʰitʃaki	right now, immediately
luntʰata	thief
mak'unku	potato meat
sajt'ata	past
qʰawqa	how (*cuánto*)
tʰahti, tʰohto	fried food made with quinoa flour
tʰata	wide
tʃʰehtʃiɾaɲa	to scorch
tʃʰojtʃu	eaves
k'ujka	intestinal worm
k'ujki	estrongilosis
p'ahpaku	nightjar, goatsucker; dance from Huancane province
q'atʃqa	fierce, rough
q'esqa	hard
laq'aqe	panting, gasping

About an equal number of forms have regular alternants: tʰanta/tʰantʰa 'old house, worn clothes', k'ankaʎi/k'ank'aʎi 'cracked skin', etc.

Although the number of forms in violation of (9e) seems rather large, it appears to be small in comparison to the number of forms of this type that would be expected if their occurrence was not restricted (but note that I have not done a statistical analysis of this effect). Furthermore, some of these forms may be spurious counterexamples: I followed a conservative strategy in eliminating those forms that I could argue were bimorphemic.

The leftness effect has few counterexamples. The two dictionaries consulted include just three relevant forms: haqt'a 'stew meat', kawtʃ'a 'how large is it', and kawq^{h}a 'from which place'.

3.2.2 Ofo

3.2.2.1 Ofo - language basics

Ofo is an extinct Siouan language formerly spoken along the lower Mississippi. All data for this language were elicited by John Swanton in 1908, first described in Swanton 1909, and fully set forth in Dorsey and Swanton 1912, which includes a dictionary of around 500 Ofo forms.

Although the existence of distinctively aspirated segments in this language has been disputed by some authors (Wolff 1950:65, Matthews 1958:13), Swanton clearly states that contrastive aspiration exists: "Probably the consonants followed by *h*, which is here very distinct, correspond to the aspirated consonants of other Siouan dialects" (Dorsey and Swanton 1912:4). More recently, Haas 1969, Rood 1979, and de Reuse 1981 have all accepted the existence of aspirated consonants in Ofo.

The consonant inventory of Ofo is given in (12). I follow de Reuse 1981 in his interpretation of Swanton's work, but I have converted de Reuse's notation to IPA transcription. De Reuse also includes, in his segment inventory, two phonemes of which he is uncertain: /b/ and /ʃ/; I omit these phonemes from consideration here.

(12)

p	t	tʃ	k
p^h	t^h	$tʃ^h$	k^h
f	s		h
f^h	s^h		
m	n		
w	l	j	

I am not aware of any languages apart from Ofo for which the segment /f^h/ has been reported; for example, none of the languages in Maddieson 1984 include this segment. Interestingly, the cooccurrence restrictions described in the next section can be taken as phonological evidence supporting Swanton's transcription of this unusual segment: the form given in (13), for instance, shows deaspiration of /k^h/ provoked by the occurrence of /f^h/ in the following syllable. Had Swanton's *fh* actually been /f/, deaspiration should not have occurred.

The vowel system of Ofo is /i/e/a/o/u/; /ĩ/ã/ũ/ and perhaps /õ/ are also present in the language.

3.2.2.2 Ofo - cooccurrence restrictions

De Reuse 1981 noted the existence of a cooccurrence restriction on aspirated segments (other than /h/) similar to that described by Grassmann's Law ("A syllable of the shape ChV loses its aspiration when it comes to occur before another syllable of the shape ChV"). A compound involving deaspiration is shown in (13).

(13)

ˈosk^{h}a 'the crane' + aˈf^{h}ã 'white' > oskaf^{h}a 'the white or American egret' (*osk^{h}af^{h}a)

Implicit in de Reuse's statement of the deaspiration rule is an intramorphemic cooccurrence restriction on aspirated elements. I have summed up the effects of this restriction in (14) (this restriction must refer to aspirated obstruents, rather than aspirated stops, due to the presence of aspirated fricatives in the language).

(14)

Heterorganic aspirated obstruents may not cooccur within morphemes: *k^{h}et^{h}i, *f^{h}ep^{h}i

By contrast, forms with more than one /h/ *are* found, as are forms with /h/ and aspirated segments. Examples are shown in (15) and (16). (I have converted Swanton's transcriptions to IPA; it is not always clear what his symbols are intended to represent, but the uncertainties do not bear on the issues discussed here. See Dorsey and Swanton 1912 for the original transcriptions.)

(15)

ahi'hi	blood
ho'he	to bellow (like a bull), to howl (like a wolf)
e'honhe	to grunt (like a pig)

(16)

in'f^{h}ihi, if^{h}i'hi	afraid, scared
in'k^{h}ehi	it is enough
'tuft^{h}ahe	to hoe
i'tʃ^{h}ohi	green, unripe

Forms with homorganic aspirated and unaspirated obstruents are also found. Examples are given in (17).

(17)

ˈtufafʰa, dufafʰa, tufafʰahi	to tear
poˈpʰuti	to swell or puff out
waˈkʰeska	drunk

In Swanton's Ofo dictionary, there are no entries containing identical aspirated stops or fricatives. This gap is discussed in the following two sections, where I suggest that the lack of such forms in the small data set available to us does not provide conclusive evidence that they did not exist in the language.

3.2.2.3 Ofo - descriptive details

There are no counterexamples to the restriction given in (14).

The occurrence of morphemes containing two /h/s leads us to ask whether or not forms with identical aspirated stops or fricatives were also able to occur. The Ofo dictionary offers no instances of such forms. However, this could be an accidental gap, arising from the small size of the dictionary. Swanton's dictionary has just 522 entries; Ofo phonotactics allow only one aspirated segment per syllable. A statistical analysis of the number of aspirated consonants and the number of words of different syllable counts suggests that although the absence of forms with heterorganic aspirated obstruents is not accidental, the lack of forms with identical aspirated obstruents could be accidental. We would expect only 1 such form in a dictionary of this size; the lack of any such forms could be due to chance. This is the line of reasoning pursued in the following section. I conclude that, based on the Ofo dictionary alone, we cannot know whether or not

identical aspirated obstruents could occur in morphemes of the language. In chapters 4 and 5, I present an analysis that indicates that the lack of Ofo forms with identical aspirated obstruents represents a genuine cooccurrence restriction of the language.

3.2.2.4 Ofo - excursus on the lack of forms with identical aspirated stops

As noted above, the Ofo dictionary includes no forms with more than one aspirated stop. The evidence I present here will show that the lack of heterorganic aspirated obstruents must indicate a cooccurrence restriction, but that the lack of forms containing identical aspirated stops could be accidental.

Dorsey and Swanton 1912 has 522 entries, comprising 1417 syllables (there are 11 entries containing vowel sequences; I treated these vowels as heterosyllabic). These 1417 syllables include a total of 201 aspirated consonants other than /h/. Aspirated consonants can appear only in onset position. Dividing the total number of aspirated consonants by the total number of syllables (201/1417) indicates that the rate of occurrence of aspirated consonants per syllable is .14.

Assume that aspirated consonants are limited to one per syllable, but occur in a freely independent fashion. There are 195 disyllables; if we let A be a syllable containing an aspirated consonant and C be its complement (i.e., a syllable without an aspirated consonant), then there are four possibilities for a disyllable: AA, AC, CA, and CC. The probability of AA is (.14)(.14) = .02. The probability of AC is (.14)(.86) = .12. The probability of CA is the same. The probability of CC is (.86)(.86) = .74.

There are 195 disyllabic entries in the dictionary, so (.14)(.14)(195) = 3.9 indicates that we could expect about 4 disyllables containing two aspirated consonants, if the aspirated consonants were occurring independently of one another (i.e., if they were not subject to a cooccurrence restriction). Similar calculations for words with three, four, five, and six syllables predict that an additional 19 forms with two or more aspirated consonants should be found, for a total of 23.

Apart from de Reuse's compounds, I did not attempt to decompose longer words into constituent morphemes. However, about four-fifths of the dictionary entries are disyllabic or trisyllabic. Even if some of the longer words are bimorphemic, this would have little effect on the results of this analysis.

Given the rate of occurrence of aspirated segments in the dictionary, and the number of forms of different syllable counts, we would expect 23 forms with two or more aspirated segments (other than /h/). Instead, we find none; a chi square goodness of fit test yields $x^2 = 24.06$. This indicates that the probability of this occurring by chance is less than .001. I conclude that the lack of forms containing two aspirated consonants is not due to chance.

Recall, however, that there are no examples of forms containing identical aspirated obstruents. A consideration of the independent occurrence of the relevant consonants suggests that this could have arisen by chance. The dictionary entries include 32 occurrences of /p^h/, 50 /t^h/, 19 /tʃ^h/, 33 /k^h/, 25 /f^h/, and 42 /s^h/; these consonants occur at rates of occurrence of .023, .035, .013, .023, .018, and .030 per syllable, respectively.

Performing calculations similar to those described above suggests that, if these consonants occurred independently, we would expect to find just one form including two identical aspirated consonants (the actual probability summed across aspirated consonants at all places of articulation is .96). The absence of one expected form would not be unusual; I conclude that we cannot assert the existence of a lexical cooccurrence restriction operating on identical aspirated consonants.

As noted earlier, the expected number of forms containing two or more aspirated consonants is 23. Subtracting 1 from this number (representing the hypothetical form containing identical aspirated consonants) will not significantly alter the results of the earlier calculations: we would expect 22 forms to contain non-identical aspirated consonants, and the likelihood of finding none – in the absence of a cooccurrence restriction – is still improbably high. The chi square goodness of fit yields $x^2 = 22.97$, indicating that the chance of this having occurred by accident is less than .001.

3.2.3 Gojri

3.2.3.1 Gojri - language basics

Gojri (also known as Gujuri, Gujer, Gujar, Gujjari, Gurjar, Gogri, Gojari, Rajasthani Gujuri, and Kashmir Gujuri) is an Indo-European:Indo-Iranian:Indo-Aryan: Central Zone:Rajasthani language. It is spoken mostly in India, but also in Afghanistan and Pakistan. A 1971 census set the number of Gojri speakers at 388,000. (This information is from Grimes 1988.) My source for this language is Sharma 1979; the segment inventory of Gojri is given in (18).

(18)

p	t	ʈ	c	k	q
p^h	t^h	$ʈ^h$	c^h	k^h	
b	d	ɖ	ɟ	g	
f	s		ʃ	x	h
	z			ɣ	
m	n	ɳ			
	l	ɭ			
	r	ɽ			
w̥			j		

The segment /h/ occurs only in initial position. The vowel system is /i/ɪ/e/ɛ/ə/a/ɔ/o/ʊ/u/; nasalization is phonemic. The language has mid, high, and low tones. Mid tones are unmarked, high tones are marked with an acute accent, and low tones are marked with a grave accent (mid tones are overwhelmingly the most common).

3.2.3.2 Gojri - cooccurrence restrictions

I am not aware of any earlier statements on cooccurrence restrictions in Gojri. Allen 1957 details the cooccurrence restrictions of Harauti, another Rajasthani language; these restrictions were mentioned in footnote 7. The Gojri restrictions are laxer than the Harauti restrictions.

The Gojri restrictions are set out in (19). These restrictions were extracted from the 1,797 words given in Sharma 1979. I am not able to specify the level at which these restrictions apply; given the data in Sharma, they could hold true of the morpheme level or of some level higher than the morpheme. In any case, the restrictions hold true of all words cited in Sharma. Because I found no instances of counterexamples to these restrictions, I do not present a section on descriptive details.

(19)

There is only one aspirated stop per word, unless the aspirated stops are identical: c^{h}ɪc^{h}ɭəp 'cobra', k^{h}ək^{h}ɽi 'a kind of cucumber'; *t^{h}ap^{h}i, *k^{h}ic^{h}a, etc.

There are at least a couple of forms illustrating the cooccurrence of /h/ with aspirated stops: hat^{h}i 'elephant', hət^h 'hand'. (A statistical analysis would be necessary to show whether or not such forms are underrepresented in the list of lexical items available to us; /h/ is a rather rare segment.) Forms with two /h/s do not occur due to the lack of forms with non-initial /h/.

Two voiceless unaspirated stops, whether identical or non-identical, do cooccur: kətɳo 'to spin', pərpato 'blunt'. Forms with two voiced stops also exist: bəd 'a piece of

land', əmdad 'help'. Finally, combinations of stops with different manners of articulation also occur: pàɲʈo 'sister's son', dəsto 'handle', kʰədɪjaɭ 'toy', bapʰən 'eyelash', pʰəʈ 'injury', kʰʊrkɲo 'to itch'.

3.2.4 Pattern Two – summary

The three languages discussed in the preceding sections – Peruvian Aymara, Ofo, and Gojri – all bear out the Pattern Two restrictions. Aspirated stops (whether voiced or voiceless) do not cooccur within roots, nor do ejectives; glottal stop also does not cooccur with itself or with ejectives. There are also no cases of ejective segments cooccurring with aspirated segments.

None of these languages have inventories that include implosives, but if such languages were found, we would expect them to include restrictions against the cooccurrence of implosives with glottal stop and with other implosives.

The Pattern Two languages illustrate the existence of the identity effect: /h/, aspirated stops, and ejectives, for example, can cooccur just in case they are identical. Ofo may not adhere to this generalization. This language is now extinct, and we cannot say, based on the available data, whether or not identical aspirated stops were able to cooccur in it. The analysis offered in chapters 4 and 5 suggests that the absence of Ofo forms containing two identical aspirated stops should be considered to represent a genuine lexical gap. In those chapters I argue that Ofo – along with Shuswap, a language discussed in 3.4.2 – constitutes an example of a second type of identity effect. The

existence of this partial, or incomplete, identity effect is predicted by the analysis presented later.

3.3 Pattern Three

The Pattern Three cooccurrence restrictions can be summed up in the statement that moderately similar segments are prohibited from cooccurring. A schematic showing which segments can and cannot cooccur within morphemes in these languages is given in Table 3.3; the languages under consideration are Bolivian Aymara ("BA"), Hausa ("H"), and Old Georgian ("OG").

	Pattern One	Attested in ...
t^h⇔ʔ	✓	BA, H
t'⇔ h	✓	BA
t^h⇔ɓ	✓	H
ɗ⇔h	✓	
h⇔ʔ	✓	BA
t^h⇔h	✓	BA
$d^{ɦ}$⇔h	✓	
t'⇔k^h	✓	BA, H, OG
t^h ⇔ $b^{ɦ}$	✓	
t^h⇔k^h	✓	BA, H, OG
$d^{ɦ}$⇔$g^{ɦ}$	✓	
t^h ⇔ $d^{ɦ}$	✓	
ʔ ⇔ɗ	*	H
ʔ⇔t'	*	BA, H
t'⇔ɓ	*	H
ɓ⇔ɗ	*	H
t'⇔k'	*	BA, H, OG
t'⇔t^h	*	BA, H, OG

h⇔h	✓	BA
ʔ⇔ʔ	✓	
t^h⇔t^h	✓	BA, H, OG
$d^{ɦ}$⇔$d^{ɦ}$	✓	
t'⇔t'	✓	BA, H, OG
ɗ⇔ɗ	✓	H

Table 3.3. Laryngeal cooccurrence restrictions of Pattern Three languages.

Bolivian Aymara and Hausa are both languages of the Pattern Three type; Old Georgian, reported on in 3.3.3, is indeterminate between Patterns Three and Four.[11]

[11] Ossetic appears to have a set of restrictions similar to those of Old Georgian. I have omitted this language from consideration because of the difficulty I had in finding accessible sources.

3.3.1 Bolivian Aymara

3.3.1.1 Bolivian Aymara - language basics

See 3.2.1.1 for basic facts on Aymara. My source for this dialect is De Lucca 1987, a dictionary of about 8000 forms. The segment inventory of the dialect I have labeled "Bolivian Aymara" is given in (20). Notice that this inventory differs from that of Peruvian Aymara (see (8)) only in that /ʃ/ is missing and /χ/ is present.

(20)

p	t	tʃ	k	q	
p'	t'	tʃ'	k'	q'	
p^h	t^h	$tʃ^h$	k^h	q^h	
	s			χ	h
	ɾ				
m	n	ɲ			
	l	ʎ			
w		j			

3.3.1.2 Bolivian Aymara - cooccurrence restrictions

The cooccurrence restrictions of Bolivian Aymara are even less taxing than those of Peruvian Aymara. Again, earlier commentary on the cooccurrence restrictions of Aymara can be found in Hardman et al. 1974, Adelaar 1986, and Landerman 1994. MacEachern 1996 discusses the ordering restrictions governing the location of aspirated and ejective stops in roots containing both types of segments, and argues for connections between these restrictions and some universal properties of segment inventories. This article is incorporated into 3.3.1.4.

The laryngeal cooccurrence restrictions of Bolivian Aymara apply at the morpheme level; they are laid out in (21).

(21)

a. There is only one ejective per morpheme, unless the ejectives are identical: k'ask'a 'acid to the taste', t'ant'a 'bread'; *t'ank'a.
b. Homorganic ejective and aspirated stops do not cooccur: t'at^{h}a.
c. Forms transcribed with initial vowels may not include ejectives: *at'a.
d. Morphemes with homorganic aspirated and plain stops, and morphemes with homorganic ejective and plain stops, are rare: *t'ata, *t^{h}alta.

The restrictions of Bolivian Aymara differ from those of Peruvian Aymara in two respects. First, Bolivian Aymara allows the cooccurrence of non-identical aspirated stops (p^{h}ut^{h}u 'hole, hollow'). Second, Bolivian Aymara allows ejective and aspirated stops to cooccur within morphemes, as long as they are heterorganic (t'alp^{h}a 'wide', p^{h}ank'a 'rubble').

3.3.1.3 Bolivian Aymara - descriptive details

De Lucca includes a handful of exceptions to the above restrictions.

There are no counterexamples to (21a) or (21b): De Lucca includes no monomorphemic forms with heterorganic ejectives, or with homorganic aspirated and ejective stops.

De Lucca includes three forms that are transcribed with initial vowels (I mention in 3.2.1.2 that I treat these forms as glottal stop-initial) but contain ejectives. These forms are amawt'a 'expert, learned person', oq'atʃaɲa 'to obstruct the way', and uktʃ'a 'stature'.

There are 16 forms with ejective or aspirated stops and homorganic voiceless unaspirated stops. Ten of the 16 forms have alternate, regular pronunciations: for

example, tʃʰiwtʃi / tʃʰiwtʃʰi 'hen pigeon' and k'akaʎi / k'ak'aʎi 'ruggedness'. The six forms without regular alternants are shown in (22).

(22)

tʰihtayaɲa	to loosen, slacken
kʰukata	away from the speaker (Omasuyos Province dialect)
t'usti	disarranged, disordered
t'utu uraqe	clay-like land
k'akaptaɲa	to crack the skin with cold
k'urkiptaɲa	to feel like crying

Finally, there is just one violation of the leftness effect. This form is quite similar to one reported for Peruvian Aymara in 3.2.1.3: kawkʰa 'where'.

I conclude that the restrictions shown in (21) are in force within the Bolivian Aymara lexicon.

3.3.1.4 Bolivian Aymara - excursus on the ordering restrictions governing roots that include both aspirated and ejective stops[12]

Bolivian Aymara exhibits an intricate pattern of wellformedness in the distribution of ejective and aspirated stops in morphemes that contain both. This pattern reveals a dispreference for /p'/ and /q'/: these sounds exist in the segment inventory of the language, but are avoided under certain conditions. This dispreference for /p'/ and /q'/ mirrors a strong crosslinguistic generalization: in languages with ejective series, the series will be defective at the labial and uvular places of articulation before it is defective at any other place of articulation. These two observations combine to suggest an implicational hierarchy such that /p'/ and /q'/ are disfavored over all other ejective stops. I appeal to phonetic motivations behind the dispreferences for /p'/ and /q'/; the analysis provided in chapters 4-6 accounts for this data.

As noted in 3.3.1.2, Bolivian Aymara observes various restrictions on the type and location of aspirated and ejective consonants that occur within morphemes. For example, if an Aymara morpheme has a single ejective or aspirated stop, then that stop will be the leftmost stop in the morpheme: k'anta 'spinning wheel', q^hatu 'market'; *kant'a, *qat^{h}u. Although ejectiveness and aspiration must occur on the leftmost stop in a morpheme, then that stop need not be morpheme-initial: sirk'u 'nerve', hajp^{h}u 'in the dark'. It is

[12] This section is condensed from MacEachern 1996.

also the case that heterorganic ejectives do not cooccur morpheme-internally (*t'ak'a), nor do homorganic aspirated and ejective stops (*q^hoq'a, *t'ant^{h}a).

Ejective and aspirated stops do cooccur morpheme-internally in the dialect of Aymara described here, but certain ordering restrictions are observed. If the initial stop in the root is a dental, palato-alveolar, or velar, then that sound will be ejective, and the second stop in the root will be aspirated. Consider the sample forms in (23) below.

(23)

t'alp^{h}a	wide
tʃ'ip^{h}a	leather net
k'ip^{h}a	said of late potatoes
k'it^{h}a	fugitive
t'ink^{h}a	tip
tʃ'ik^{h}i	clever
t'aq^{h}e	affliction
tʃ'oʌq^{h}e	solid

De Lucca includes 58 roots that fall into this category. Not all possible place of articulation combinations are attested, however, and some of these gaps do not appear to be accidental. For example, the lack of roots of form /t'...tʃ^h/, /tʃ'...t^h/, /k'...q^h/ and /q^h...k'/ may follow from prohibitions on the cooccurrence of similar, but non-identical coronal and back lingual articulations. This hypothesis is supported by the lack of Aymara roots of form /t^h...tʃ^h/, /tʃ^h...t^h/, /k^h...q^h/, and /q^h...k^h/. I do not pursue this here.

The other side of the ordering restrictions concerns forms with an aspirated and an ejective stop in which the initial stop is labial or uvular, and the second stop is at an

intermediate (i.e., dental, palato-alveolar, or velar) place of articulation. In this case, the first stop in the root is aspirated, and the second is ejective. De Lucca includes 34 roots of this type (a few exceptions exist; these are discussed below); several are shown in (24).

(24)

p^hant'a	black coat
p^hitʃ'i	coat pin
p^hank'a	rubble
q^hot'a	resin of some small plants
q^hatʃ'u	fodder

The restrictions described above entail that, if the places of articulation of the stops are known, and if the number of aspirated and ejective features within the morpheme is known, then the location of the laryngeal features is predictable.

There is one subset of data for which the statement above does not hold true. De Lucca includes a few forms with one ejective and one aspirated stop, one of which is labial and the other of which is uvular. There are four possible arrangements of two stops and two laryngeal features; these are shown in (25).

(25)

q' ... p^h
q^h ... p'

p' ... q^h
p^h ... q'

Of the four patterns given in (25), only two are attested. An exhaustive list of these forms is given in (26).

(26)

q'ap^{h}i	fragrance
q'ap^{h}a/q'apa	active, diligent, agile
q^{h}op'aki	meal with meat or fat (Altiplano Central)
q^{h}op'i	potter (Altiplano Norte)

Notice that the roots of form /q^h ... p'/ are singled out by De Lucca as being dialectal variants, while the roots of form /q' ... p^h/ stand as representatives of the main dialect. There are no roots of form /p'... q^h/ or /p^h...q'/; this gap might be accidental. I will return to these facts in 6.2.

Regarding the exceptions mentioned above, there are six forms with initial ejective uvulars that are followed by aspirated stops at intermediate (i.e., non-labial, non-uvular) places of articulation. The account given above leads us to expect these forms to involve aspirated uvular stops followed by ejectives. However, four of these exceptional forms have alternate, regular pronunciations in De Lucca or in Miranda 1970 (another Aymara dictionary published in Bolivia). The two residual forms are q'ajt^{h}i 'beating of waves' (Manco Capaj dialect) and q'atʃ^{h}i 'edge'; the expected forms are *q^{h}ajt'i and q^{h}atʃ'i. I have no explanation for these roots; they do not necessarily indicate a milder dispreference for /q'/ as opposed to /p'/, given that /q'/ occurs about twice as frequently as /p'/ in this language. In the rest of this study, I disregard these two forms.

As noted earlier, the Bolivian Aymara dispreference for /p'/ and /q'/ mirrors a strong crosslinguistic generalization: in languages with ejective series, the series will be

defective at the labial and uvular places of articulation before it is defective at any other place of articulation. This statement, established through reference to several compilations of segment inventories from many languages, builds on the observations of Greenberg 1970 and Fordyce 1980. Those authors pointed out that the presence of /p'/ in an inventory implies the presence of ejectives farther back in the oral cavity.[13]

The investigation of defective ejective series is complicated by the crosslinguistic tendency to have ejectives only at places of articulation where another stop series is instantiated (Greenberg 1970; Fordyce 1980 proposes the "ejective-to-plain" hierarchy to describe the tendency of languages to have ejectives only where they also have plain voiceless stops). Assuming that this characteristic is accounted for by other features of the grammar, we might expect a crosslinguistic dispreference for bilabial and uvular ejectives to reveal itself in this way: if a language lacks an ejective at some place of articulation where another stop series is instantiated, then the missing ejectives will be bilabial and/or uvular; if both /p'/ and /q'/ are missing, then ejectives at other places of articulation may also be absent.

This prediction is borne out by a review of the data in Maddieson 1984 and Ruhlen 1975, 1976.[14] Of the nearly 100 languages in Ruhlen and the roughly 35

---

[13] The work of Campbell 1973 is also important in connection with this; Campbell pointed out that many Mayan languages have uvular implosives despite having ejectives at forward places of articulation. This observation established that uvular ejectives were not bound to occur wherever labial ejectives occurred.

[14] Greenberg 1970 states that Amharic and some other Semitic Ethiopian languages have /k/ and /q'/ (but not /q/), rather than /k/ and /k'/. Such a language would constitute a solid counterexample to my claim. However, Greenberg does not cite his sources for this

languages in Maddieson with either /p'/ or /q'/, or both (of course, most of Maddieson's languages also appear in Ruhlen), only a handful are described as lacking or having marginally-attested ejective stops at places of articulation present in another stop series of the language (dental, alveolar, palatal, velar, or labiovelar, with or without a secondary articulation). I will argue that none of these languages present a robust counterexample to my claim. By contrast, several Na-Dene languages (Chipewyan, Haida, etc.), several Mayan languages (Aguacatec, Huastec, etc.), several Northeast Caucasian languages (Andi, Bagvali, etc.), West Circassian, Mazahua, Gununa-Kena, and Ossetic all include ejectives at intermediate places of articulation, but are missing either /p'/ or /q'/ or both of these segments, although labial and/or uvular stops of other series are present.

The nearest exceptions to the generalization I have given are Huambisa (an Andean language of the Jivaro family), Chol (Mayan), Tewa (Tanoan), Osage (Siouan), and Berta (Nilo-Saharan:Chari-Nile). I will now explain why none of these five languages constitutes a good counterexample to the claim made above.

Huambisa (Beasley and Pike 1957) is listed as having only /p'/ in the ejective series, although /t/ and /k/ are also present in the language. However, /p'/ is found only as an expressive interjection among men: if this segment is not really a phoneme of the language, then the lack of /t'/ and /k'/ does not violate the generalization invoked here.

---

observation; the sources I have checked do not indicate that this is true of any Semitic Ethiopian languages.

In three of the remaining four languages, one of the relevant sounds is extremely marginal. In Chol and Tewa, an ejective at an intermediate place of articulation is missing, but the matching non-glottalic stop is quite rare. Chol (Warkentin and Brend 1974) has two full series of aspirated and ejective stops, except that it lacks /t'/, although it has /t^h/. However, /t^h/ is very rare. Similarly, Tewa (Hoijer and Dozier 1949) has /tj̇/ but lacks matching /tj̇'/; however, the authors report that /tj̇/ is beyond rare: it was only found in one morpheme (tj̇u 'younger sibling'). The rareness of the relevant plosives (/t^h/ in Chol and /tj̇/ in Tewa) suggests two possibilities: either the plosives (stops produced with the pulmonic airstream mechanism) are not fully integrated into the segment inventories of the languages under consideration (in which case the lack of a matching ejective is not surprising), or the lack of matching ejectives is accidental. The second explanation rests upon the crosslinguistic tendency of ejectives to be much less frequent than the corresponding plosives.

Osage (Wolff 1952) is said to have /p'/ and /k'/ but lack /t'/, although it has /t/. However, Wolff notes that /k'/ is reconstructed to Proto-Siouan */kq/, while Proto-Siouan */pq/ and */tq/ normally become Osage /p/ and /t/, respectively. This observation suggests that /p'/ might be marginal. In fact, the only /p'/ mentioned in the article is in tap'ok'e, which is labeled as an allomorph of tap^hok^he ('he hit(s) it'). This suggests that /p'/ is extremely rare, and perhaps not even distinctive. If /p'/ is missing from the segment inventory, of course, then the absence of /t'/ is not remarkable.

Finally, for Berta, Triulzi et al. 1976 offer the stop inventory shown in (27) (the affricate /dʒ/ is also present).

(27)

p'		k'
b	d	g

According to the claim made here, we would expect unattested /t'/ to also be present in the segment inventory. However, the authors note that /p'/ and /k'/ are "weakly glottalized and often approach [p, k] in realization". The lack of a voiceless plosive series in Berta means that there is no need to maintain a contrast between /k'/ and /k/ or /k^h/, for example. Furthermore, "[t]he fricative /θ/ seems to fill the position of the missing t' in the stop series" (Triulzi et al. 1976:520). It is also suggestive that there appears to be a great deal of dialectal variation in the ejective series of Berta. Apart from Triulzi et al. 1976, none of my sources (Reidhead 1946, Cerulli 1947, Andersen 1993) list /p'/ as a segment of the language.[15]

Some functional bases for the dispreference for /p'/ are described below; weakly glottalized segments will be less subject to these dispreferences. In the absence of phonetic data on the Berta dialect group, I do not consider it a good counterexample to the generalization argued for here.

[15] One other language may be relevant here. The Lezghian language Agul (Kumykh) has an extensive stop series including /p'/t'/k'/q'/q^w'/, but may be missing /k^w'/, although /k^w/ is attested. I have not been able to gather complete information on this language.

I will now consider the possible bases of the dispreferences for /p'/ and /q'/. Various scholars (Haudricourt 1950, Wang 1968, Greenberg 1970) have noted that ejectives tend to occur at back rather than forward places of articulation; Fordyce 1980 showed that this holds true over a large language sample, and Maddieson 1984 established that it was primarily bilabial ejectives that were disfavored. Three rationales for the dispreference for /p'/ have been suggested by other scholars. First, the bursts of bilabial stops are relatively weak (Zue 1980) due to the lack of a downstream resonating chamber (Kawasaki 1982). Second, ejectives involve compression of the air in the supraglottal chamber. Because bilabial constrictions involve the largest supraglottal chamber, the compressive effect will correspondingly be least significant in /p'/ (various authors beginning with Wang 1968 have noted this, but see especially Javkin 1977). Third, a strong compressive effect will be difficult to achieve in bilabial stops because they expose the greatest amount of yielding cheek wall surface (Kingston 1985, who follows Ohala and Riordan 1979 on the significance of passive compliance of the vocal tract walls). All three of these explanations rely on the assumption that less salient sounds will be disfavored crosslinguistically.

The factors just cited for the dispreference of /p'/ would seem to indicate that /q'/ should be a preferred articulation. There is a downstream resonating chamber, and compression of the air in the oral cavity should be easy to achieve for two reasons: first because the oral chamber is small during uvular stop production, and second because it is walled with tenser material than the cheeks. The Aymara data, however, suggest that

uvular ejectives are disfavored. It may be the case that maintaining a uvular seal during ejective production is difficult when the supraglottal chamber is small and when both of the articulators involved (the tongue dorsum and the uvula and lower velum) are soft. The elevated oral pressure characteristic of ejectives might tend to make the uvular closure slip, resulting in a uvular affricate or fricative.

I conclude that a crosslinguistic study of segment inventories reveals no strong counterexamples to the generalization stated above: no language will have /p'/ or /q'/ unless it also has ejectives at all other places of articulation for which stops are attested. I take this typological generalization as evidence of the markedness of labial and uvular ejectives. Aymara is a language in which labial and uvular ejectives are present; however, the data presented earlier illustrates the emergence of the unmarked (McCarthy and Prince 1994) – or rather, less marked – in a small corner of the lexicon of this language. In Aymara, ejective features generally appear as near the beginnings of morphemes as possible. However, this rule is violated just in case the following three statements hold true: (1) assigning ejective features to the leftmost stop in the morpheme would create an ejective labial or uvular segment, (2) another host for the ejective feature can be found, and (3) satisfaction of the leftness effect, which requires aspiration and ejective features to be as near the beginning of the word as possible, can be accomplished by other means (i.e., by placing aspiration before ejectiveness). In just these instances, the aspiration feature will precede the ejective feature, apparently in order to avoid the creation of ejective labial and uvular segments. The typological generalization and the

ordering facts find a unified explanation in the statement that /p'/ and /q'/ are marked (i.e., they are disfavored over all other ejective stops). See chapters 4-6 for analysis.

3.3.2 Hausa[16]

3.3.2.1 Hausa - language basics

Hausa is Afro-Asiatic:West Chadic (Plateau Sahel); it is spoken as a first language by about 5,700,000 people (and as a second language by many millions more), mostly in Nigeria and Niger (Grimes 1988, citing a 1952 survey).

Hausa contrasts voiced, voiceless, and glottalic obstruents. The glottalic stop series has implosive stops at anterior places of articulation, and ejective velar stops. The consonant inventory is given in (28) (Kraft and Kraft 1973; palatalized and labialized velar series added). Note that I transcribe the voiceless plosive series as aspirated. Although these segments are not generally referred to as aspirated, Ladefoged 1964 states that the average period of voice onset time for these segments is 35-45 ms.

[16] I am grateful to Russell Schuh and Alhaji Gimba for help with the Hausa data. All errors, of course, are my own.

(28)

b	d		dʒ		g	g^j	g^w
ɓ	ɗ				k'	k^{j}'	k^{w}'
	t^h		$tʃ^h$		k^h	k^{jh}	k^{wh}
	ts'						
				ɥ			ʔ
	z						
ɸ	s		ʃ				h
	ɾ	r					
	l						
w				ɥ			
m	n						

The labial fricative /ɸ/ is restricted to coda position. The laryngealized labio-palatal glide (/ɥ/) is rare, and in fact all instances of this sound appear to derive from a single root, ɗi 'child' (Russell Schuh, personal communication). Glottal stop appears initially in all forms that are standardly written as vowel-initial; I have added glottal stop to these forms where they are cited below. Glottal stop and /h/ are both rare in word-medial position; most forms that include these segments in medial position are Arabic borrowings (Parsons 1970).

The vowel inventory of Hausa includes long and short vowels of five qualities: /i/e/a/o/u/. The language has high, low, and high-falling tones.

3.3.2.2 Hausa - cooccurrence restrictions

The Hausa cooccurrence restrictions are given in (29). Parsons 1970 was the first to comment on these restrictions; Hopper 1973 highlighted the crosslinguistic significance of restriction (29a). These statements were checked against Abraham 1962, a dictionary of around 35,000 forms. All data comes from this dictionary; low tones are marked with grave accents.

(29)

a. Heterorganic glottalic segments (whether ejective or implosive) do not cooccur within roots: *ɓaɗa, *ɓak'a, etc.
b. Glottal stop does not cooccur with glottalic segments, whether those segments are ejective or implosive: *ʔaɓa, *ʔak'a, etc.
c. Glottalic segments, whether ejective or implosive, tend not to cooccur with homorganic voiceless aspirated stops: *ɗatʰa, *k'akʰa, etc.

Other segment combinations do occur. For example, /h/ occurs with glottalic segments, aspirated stops, and voiced stops: hàɗa 'to join', hàk'uɽà 'is patient', huutʰà 'to rest', hadirì 'storm'. Aspirated stops also occur with other aspirated stops, and with heterorganic ejectives, implosives, and voiced stops, as well as with glottal stop: kʰeetʰà 'to rip', k'aatʰò 'huge', ɓatʰa 'spoiled', kʰago 'round, thatched house', ʔiikʰò 'power, control'. Finally, voiced stops also occur with other voiced stops, with glottal stop, and with implosives: buga 'thrashed', ʔagùgù 'plant used as remedy for tapeworm', ɗigò 'drip, drop'.

Roots may also contain identical aspirated stops (tʰootʰùuwa 'pulp of guords', tʃʰaatʃʰa 'gambling'), identical glottalic segments (k'ook'ì 'wife of mythical spider', ɗaaɗà 'sharpened edge of tool', ɓuuɓà 'bargain', ts'uuts'à 'worm'), and identical voiced stops (gigè 'place').

There are also cases of dictionary entries containing two /h/s and two glottal stops: huhà 'baobab tree', hoho 'shells of beans, groundnuts, etc.', ʔaaʔà 'no', ʔìiʔìinaa 'custom'. Although some of these forms manifest identical vowel qualities in the first two syllables, none are recognized as involving productive reduplication. Recall that most forms that include two /h/s or two /ʔ/s are loanwords.

3.3.2.3 Hausa - descriptive details

The only exceptions to restriction (29a) are a pair of Arabic loanwords: ɗaaɓà 'printed' and ɗaaɓì/ɗaɓʔì 'act of printing'.

Abraham includes around 40 apparent counterexamples to restriction (29b). The overwhelming majority of these forms are noted as being Arabic loanwords; others are bimorphemic, or ideophonic expressions. Examples of some Arabic borrowings are shown in (30).

(30)

ɗaaʔù	(the letter) d
ʔàk'allà	but few
k'aaʔidà	law
ʔitt^{h}ìfaak'aaŋ	for sure

There are a total of five forms that may be genuine counterexamples to (29b); I could not find any evidence that these forms, which contain both glottal stop and /ɗ/, are polymorphemic or non-native. These words are shown in (31).

(31)

ʔàgàɗaa	beggar-musicians
ʔaɗɗè	presumptuousness
ʔagùɗii	hip-joint
ʔak^{h}àrɗàa	a food of uncooked flour with rama-leaves and water
ʔak^{h}azauɗà	long-headed, long-shafted arrow

The restriction listed in (29c) has one general class of exceptions: those of form t^{h} ⇨ɗ, where the right-pointing arrow indicates that the /t^{h}/ precedes the /ɗ/. Sample forms are shown in (32).

(32)

t^{h}abɗeeɗè	huge
t^{h}aaɗì	chatting
t^{h}anɗà	licked
t^{h}itt^{h}iɗa	gave push from below
t^{h}urbùɗaa	thrust into hot ashes to cook

Regarding the other combinations of homorganic glottalized and voiceless aspirated stops, we find no non-ideophonic, monomorphemic cases involving velars. We find just three cases of ɗ ⇨ t^h; these three forms are given in (33).

(33)

ɗaat^{h}à	the tomato-like gaut^{h}aa
ɗat^{h}ò	gravy
ɗeet^{h}àraa	crossed over

The analysis presented in chapter 5 accounts for the cooccurrence restrictions set out in (29).

3.3.3 Old Georgian

3.3.3.1 Old Georgian - language basics

Old Georgian (a South Caucasian language) was spoken from the 4th century AD onwards for several hundred years (Fahnrich 1991).

The consonant inventory of Old Georgian is given in (34) (Fahnrich 1991).

(34)

b	d	dz	dʒ	g		
p^h	t^h	ts^h	$tʃ^h$	k^h	q^h	
p'	t'	ts'	tʃ'	k'	q'	
v		z	ʒ		ʁ	
		s	ʃ		χ	(h)
m	n					
	l					
	r					
w			j			

The segment /h/ appears only as a pronominal prefix and in borrowings (Vogt 1988).

The vowel inventory of Old Georgian was /a/e/i/o/u/. Native roots tend to be monosyllabic.

3.3.3.2 Old Georgian - cooccurrence restrictions

Hopper 1973 stated that heterorganic ejectives did not cooccur in Old Georgian. My study of the Molitor 1952 lexicon indicates that homorganic ejectives and aspirated stops also did not cooccur within morphemes.

These restrictions are set out in (35).

(35)

a. Heterorganic ejectives do not cooccur: *t'ap', etc.
b. Homorganic aspirated and ejective stops do not cooccur: *$t'at^h$, *p^hap', etc.

In searching for data on the cooccurrence restrictions of Old Georgian, I was hampered by the lack of a lexicon broken down by morphemes. Fahnrich states that most

Old Georgian roots were of form CVC, and that the vowel often disappeared. That the lexicon I used – Molitor 1952, which includes about 3000 forms – is not really appropriate for my purposes can be judged by the length of many of the forms I cite. Most polysyllabic forms are polymorphemic.

The statements in (35) depend on the following two pieces of evidence: (1) Hopper's statement that heterorganic ejectives do not cooccur within native roots, and (2) the paucity of forms in Molitor – of any length – that include homorganic aspirated and ejective stops. If there are so few *words* containing this combination of segments (I found just five, two of which are clearly borrowings; this is discussed further in 3.3.3.3), then the restriction against their cooccurrence within roots seems certain.

As noted above, the segment /h/ is found only as a pronominal prefix in the native vocabulary. Therefore, we do not expect to find it cooccurring with aspirated stops in native roots. The laryngeal fricative does cooccur with aspirated stops in a couple of borrowings: hrak^{h}el 'Rachel', hrup^{h}e: 'Rufus'.

Aspirated stops also cooccur, whether heterorganic or homorganic: possible examples include t^{h}ap^{h}li 'honey', ag:ʃp^{h}ot^{h}ebaj 'disconcert', mok^{h}alak^{h}e: 'citizen'. Heterorganic ejective and aspirated stops also cooccur: daq'op^{h}a 'close', mok'ap^{h}aj 'hack off', mts'rap^{h}l 'hurried'. Identical ejectives also cooccur: ts'usts'ut^h 'often', tʃ'urtʃ'eli 'equipment' (most forms with identical ejectives involve what Fahnrich labels the pre-alveolar and post-alveolar affricate series). The rarity of /h/, of course, does not give us the opportunity to observe native morphemes with two /h/s.

3.3.3.3 Old Georgian - descriptive details

Hopper confined his remarks to Old Georgian, so we might guess that cooccurrence restrictions have disappeared from Modern Georgian. Indeed, Cherkesi 1950, a dictionary of Modern Georgian, includes many entries with heterorganic ejectives in close proximity; homorganic aspirated and ejective consonants also appear to cooccur within morphemes.

Because I was unable to break words down into morphemes, I do not attempt to present a list of the entries within Molitor that contain counterexamples to the restriction listed in (35a). However, because the restriction against homorganic ejective and aspirated stops has not, to my knowledge, been stated in the past, a list of the entries in Molitor that contain this combination of segments is shown in (36). The two obvious loanwords are provided in the second group of items. The first three forms listed may also turn out to be loanwords or polymorphemic.

(36)

helis-q'oq^{h}aj	undertake
sameots daat^h hut^{h}met'i	sixty and fifteen
t'virt^{h}i	last
p^{h}ilip'e:	Philippus
t'alit^{h}a	Talitha

3.4 Pattern Three – summary

Bolivian Aymara and Hausa both bear out the Pattern Three restrictions. As can be seen in Table 3.3, the absence of implosives and glottal stop in Old Georgian means that the four cells that distinguish Pattern Three from Pattern Four (i.e., ʔ ⇔ɗ, ʔ⇔t', t'⇔ɓ, and ɓ⇔ɗ) cannot be filled in: The cooccurrence restrictions of Old Georgian are thus indeterminate between these two degrees of strictness.

In Pattern Three languages, no glottalic segments, whether ejectives, implosives, or glottal stop, can cooccur within morphemes, unless those segments are identical. Thus, Pattern Three languages illustrate the existence of the identity effect. Furthermore, aspirated stops do not cooccur with homorganic ejectives in these languages.

3.4 Pattern Four

Pattern Four cooccurrence restrictions are the laxest laryngeal cooccurrence restrictions in existence. In these languages, only extremely similar segments fail to cooccur. This pattern is instantiated by several Mayan languages, including Tzutujil. Shuswap is also a Pattern Four language, although it has a different identity effect from that found in Tzutujil. Table 3.4 shows which segments are and are not allowed to cooccur in these two languages. Because this chart does not generalize over languages, only the relevant subset of cells is filled in.

	Tzutujil	Shuswap
t^h⇔ʔ		
t'⇔ h	✓	
t^h ⇔ɓ		
ɗ⇔h	✓	
h⇔ʔ	✓	
t^h ⇔h		
$d^{ɦ}$⇔h		
t'⇔k^h		
t^h ⇔$g^{ɦ}$		
t^h ⇔k^h		
$d^{ɦ}$⇔$g^{ɦ}$		
t^h ⇔ $d^{ɦ}$		
ʔ⇔ɗ	✓	
ʔ⇔t'	✓	✓
t'⇔ɓ	✓	
ɓ⇔ɗ		
t'⇔k'	*	*
t'⇔ t^h		
h⇔h	✓	
ʔ⇔ʔ	✓	✓
t^h ⇔ t^h		
$d^{ɦ}$⇔$d^{ɦ}$		
t'⇔t'	✓	*
ɗ⇔ɗ		

Table 3.4. Laryngeal cooccurrence restrictions of Pattern Four languages.

3.4.1 Tzutujil

3.4.1.1 Tzutujil - language basics

The Mayan family includes about thirty languages and is spoken by several million people in Mexico, Guatemala, Belize, and Honduras. Tzutujil belongs to the

Eastern:Quichean branch of Mayan (Kaufman 1974), and is spoken by about 50,000 people in midwestern Guatemala; the dialect discussed here is that of San Juan La Laguna (Dayley 1985).

The consonant inventory of Tzutujil (from Dayley 1985) is given in (37).

(37)

p	t	ts	tʃ	k	q	
ɓ	ɗ	ts'	tʃ'	k'	q'	ʔ
	s		ʃ			h
m	n					
		l				
		r				
w			j			
(b)	(d)			(g)		

The parenthesized stops (/b/d/g/) occur only in recent Spanish loans.

With respect to the transcription of the voiceless plosive series, Dayley states that consonants are aspirated in syllable-final position, but are unaspirated prevocalically. I have recorded this series as unaspirated, and will transcribe these segments as such even in final position. In doing so, I follow Lombardi 1991 in her proposal (made with reference to another Mayan language, Tojolabal) that statements like Dayley's describe plain, unaspirated stops that are audibly released in final and pre-consonantal position. (Note, however, that this point is not critical to my analysis: the analysis presented in chapters 4 and 5 predicts the same sets of occurring and non-occurring forms, whether the final voiceless plosives are understood to be aspirated or unaspirated.)

Several points of allophonic variation are crucial to an understanding of the cooccurrence restrictions discussed here. Dayley reports that /ɓ/ and /ɗ/ are realized as [p'] and [t'], respectively, in syllable-final position. The segment /q'/ is ejective except prevocalically, where it is optionally realized as a voiced implosive (i.e. [ɠ]). The fricative /h/ is [h] syllable-internally (i.e., when it appears immediately before a coda consonant); elsewhere it is [χ].

Glottal stop appears contrastively in word-medial and word-final position; it is not contrastive in initial position. Dayley reports that most monosyllables that do not begin with some other consonant have initial glottal stop; in polysyllables that do not begin with some other consonant, initial glottal stop is optional.

Tzutujil includes long and short vowels of five qualities: /i/u/e/o/a/.

Most roots are of form CVC, CV:C, CVʔC, or CVhC, but there is also "a large number" of bisyllabic roots.

Stress occurs on the final syllable of native words.

3.4.1.2 Tzutujil - cooccurrence restrictions

Many authors have described the existence of cooccurrence restrictions operating over glottalized segments in the Mayan languages. To my knowledge, the earliest such statement is found in Weathers 1947, writing on Tsotsil.

Dayley 1985 states that in Tzutujil, non-identical glottalized consonants do not cooccur within syllables, unless one of them is /ɓ/. Note that Dayley does not exempt the

other Tzutujil implosive, /ɗ/, from this restriction. However, /ɗ/ is extremely rare in Tzutujil (and indeed within the Quichean languages more generally); although the grammar provides no examples of /ɗ/ cooccurring with phonemic ejectives, this might be accidental. If this is so, then both implosives should be exempted from the cooccurrence restriction, and the descriptive statement would be that non-identical *ejectives* do not occur within the relevant domain.

With respect to Dayley's statement that the domain of the restriction is the syllable, I note that all observers of laryngeal cooccurrence restrictions in the Mayan languages have noted either that they apply only to native roots (which are typically monosyllabic), or that they apply only at the syllable level. Unfortunately, although these two statements are not equivalent, I do not have enough data on the necessary items (loanwords, polymorphemic monosyllables, and polysyllabic native morphemes) to check the level at which these restrictions apply. In the absence of more information, then, I accept Dayley's statement that these restrictions apply at the syllable level.

The Tzutujil restrictions are summarized in (38).

(38)

a. There is only one ejective per syllable, unless the ejectives are identical: tʃ'ihtʃ' 'metal, car', q'iiq' 'north wind'; *tʃ'iiq'.
b. Syllables with homorganic implosive or ejective and plain stops do not occur: *ɓaap, *taat'.

In Tzutujil, unlike Pattern Three languages, glottal stop can cooccur with ejectives (tʃ'uuʔ 'fish', ʔak' 'chicken', tʃ'aʔk 'a boil') and with implosives (ɓoʔχ 'cotton', ɓiiʔ 'name', ʔooɓ 'phlegm, cough'). I found no examples of forms containing both glottal stop and /ɗ/, presumably due to the rarity of /ɗ/. Roots with an implosive and an ejective cooccurring within one syllable are also present: ɓatz' 'thread', χuɓiit' (/huɓiiɗ/) 'a tear'. The laryngeal fricative occurs with a variety of segments, as do plain voiceless stops. Finally, we also find roots with tautosyllabic glottal stops: χaʔeeʔ 'they', kaʔiʔ 'two', ʔoʔ 'poo-poo'.

Although I did find Tzutujil *roots* with two implosives, there are no syllables with two implosives. This is because implosives become ejectives when they occur in syllable-final position. So although we find, for example, the root ɗeɓ-, appearing in ɗeɓeli 'thick (of liquid)', it would surface as ɗep'- when followed by a consonant. These distributional restrictions prevent two implosives from occurring within a single syllable.

Table 3.4 indicates that two /h/s can cooccur within Tzutujil syllables. Actually, the fricative in question is realized as [h] before coda consonants, and as [χ] elsewhere. Thus, although we find forms such as oχoχ 'we', there are – strictly speaking – no syllables containing two [h]s.

3.4.1.3 Tzutujil - descriptive details

Dayley 1985 includes just one counterexample to the restriction given in (38a): tʃ'aaɓ 'reflection, arrow' will appear as [tʃ'aap'] when it occurs without a suffix.

We find three counterexamples to restriction (38b): kik' 'blood', q'eq ~ q'aq 'black', and [ɗoot'] (/ɗood/) 'snail'.

I will disregard these exceptions, and conclude that the restrictions given in (38) hold true of Tzutujil.

3.4.2 Shuswap

3.4.2.1 Shuswap - language basics

The Salishan languages, of which there are over 20, used to be spoken across a great stretch of territory including parts of Washington, northern Idaho, southern British Columbia, and northern Oregon (Thompson and Thompson 1985). Shuswap, spoken by about 500 people in southern British Columbia (Grimes 1988), is the northernmost of the Northern Interior Salish languages (Kuipers 1974).

My source for this language is Kuipers 1974, which covers the dialects of Canim Lake and Alkali Lake. This grammar includes a dictionary of about 2000 morphemes. The segment inventory for Shuswap is given in (39).

(39)

p'	t̪l' ~ t'		ts'		k'	k^{w}'	q'	q^{w}'	
p	t		ts		k	k^{w}	q	q^{w}	ʔ
		ɬ	s		x	x^{w}	χ	$χ^{w}$	(h)
m	n	l		j	ɣ	w	ʕ	$ʕ^{w}$	
$m^{ʔ}$	$n^{ʔ}$	$l^{ʔ}$		$j^{ʔ}$	$ɣ^{ʔ}$	$w^{ʔ}$	$(ʕ^{ʔ})$	$ʕ^{wʔ}$	

The parenthesized segments are very rare; /h/ is limited to initial position. One segment is sometimes realized as [t'], and at other times as [t̪l']. What Kuipers refers to as the "dental-palatal" series, transcribed here /ts/ts'/s/, is said to vary in pronunciation between dental and alveopalatal.

The glottalized sonorants (shown in the last row of (39)) are preglottalized in prevocalic position and postglottalized elsewhere.

The Shuswap vowel inventory includes /i/u/e/o/a/ʌ/; [ə] appears in unstressed syllables.

Most roots are of form CVC or CC; CVCC and CCVC are also found. Other forms exist but are much less common. The location of stress is not predictable. The language involves extensive affixation, mostly suffixal.

3.4.2.2 Shuswap - cooccurrence restrictions

Thompson and Thompson 1985 refer to a variety of dissimilation effects operating in several Salishan languages, including Shuswap; they cite Eric Hamp as an early observer of these effects. The only Salishan language I will address here is Shuswap, but certainly the form of laryngeal cooccurrence restrictions in the other Salishan languages would be a fruitful topic for exploration (in particular, the presence or absence of the identity effect in the different languages).

Kuipers 1974 made the following observations on Shuswap cooccurrence restrictions ('K' indicates an obstruent; 'R' is a resonant):

> If a root has the shape K_1VK_2, K_1VRK_2, K_1RVK_2, and K_2 is glottalized, then K_1 is never glottalized. In any type of reduplication, the first occurrence of a reduplicated obstruent is never glottalized. Thus ...p'... is reduplicated ...p...:p'..., and, in particular, t' [t̪l'/t'] is reduplicated by t [t].
>
> The glottal stop as K_2 does not prevent glottalization in K_1, e.g. <u>c'iʔ</u> 'deer', <u>k'ʔem</u> 'put', etc.
>
> Kuipers 1974:23 (symbols updated and underscoring added)

By contrast, forms containing ejectives and glottalized sounds *do* exist: <u>s-p'uqʷiʔ</u> 'silvery shrub with small white berries', <u>p'ələnʔ</u> 'bark of tree'.

Kuipers' dictionary includes no examples of any roots or suffixes containing more than one ejective, whether those ejectives are homorganic or heterorganic. Accordingly, I have restated the Shuswap laryngeal cooccurrence restriction as shown in (40).

(40)

Ejectives, whether identical or not, do not cooccur within roots: *p'ats', *p'ap', etc.

Unlike languages of Patterns One, Two, and Three, Shuswap ejectives *can* occur with glottal stops: t'eʔ 'look for', p'eʔ-m 'to pack (a child) on one's back', qwʔʔ 'to tie on'. Unlike languages with the complete version of the identity effect, Shuswap allows ejectives to occur with homorganic plain stops: t'emt 'three year old beaver', t'uxʷt 'to fly', lqʷ'uqʷ 'chicken'.

Although Shuswap does not allow identical ejectives to cooccur, it does allow glottal stops to cooccur: ʕʔ iʔl 'to lose (a contest, law case, etc.)', xʔuʔcn-s 'to agree with', ʔuʔse 'egg'. I call this the incomplete, or partial version of the identity effect.

Shuswap makes use of several types of reduplication. When an ejective is reduplicated, the ejective in the reduplicant is deglottalized. For example, tətet'χ-t 'taller' (from t'eχ-t 'tall'), cqeq'łp 'small tree' (from cq'ełp 'fir, tree'). This indicates that the domain of laryngeal cooccurrence restrictions in Shuswap is larger than the morpheme level. As Thompson and Thompson note, the cooccurrence restriction does not hold across roots and suffixes: t-k'm-its'eʔ 'surface, bark of root', for example, shows a root containing an ejective followed by a suffix with an ejective. The root does not undergo deglottalization. From this we gather that the Shuswap laryngeal

cooccurrence restrictions operate over morphemes, and over roots plus reduplicated elements.

Now consider forms with reduplicated glottal stops. For example, q^{w}'ʔeʔχ^{w}-k 'I am skinny' (from q^{w}'ʔeχ^{w} 'skinny'), tətʔiʔk^{w}-m 'to glow, be red-hot' (from tʔik^{w} 'fire'), x-k'eʔeʔ-t '(be) jailed' (cf. k'ʔem 'to put', x-k'-nt-es 'he puts it in'). The glottal stop in the reduplicant remains; this is what we would expect, since glottal stops are allowed to cooccur through the partial identity effect.

The analysis presented in chapters 4 and 5 accounts for the existence of forms containing two glottal stops, and the absence of forms containing identical ejectives.

3.4.3 Pattern Four – summary

Tzutujil and Shuswap are the languages discussed in the preceding sections. These languages bear out the only Pattern Four restriction that is relevant, given their segment inventories: heterorganic ejectives do not cooccur. If these languages had aspirated stops, we might also expect to find a lack of morphemes including homorganic aspirated and ejective stops.

Tzutujil and Shuswap instantiate different versions of the identity effect. The Tzutujil identity effect is that which is familiar from Peruvian Aymara, Gojri, Bolivian Aymara, Hausa, and Old Georgian: similar elements – for example, two ejectives – are allowed to cooccur just in case they are identical. The Shuswap identity effect, which we might call the partial identity effect, is more complicated. In that language, two glottal

stops may cooccur within morphemes, but two ejectives may not. Shuswap also allows the cooccurrence of ejectives with otherwise identical voiceless unaspirated stops; languages with the standard, or complete identity effect do not. Earlier in this chapter, I suggested that Ofo is another language with the partial identity effect. These facts are explained by the analysis presented later in this study.

3.5 Summary

Table 3.5 summarizes the laryngeal cooccurrence restrictions examined in the languages discussed here. A key to the language abbreviations is given in (41).

(41)

CQ	Cuzco Quechua
SB	Souletin Basque
Skt	Sanskrit
PA	Peruvian Aymara
Ofo	Ofo
Goj	Gojri
BA	Bolivian Aymara
H	Hausa
OG	Old Georgian
Tzu	Tzutujil
Shu	Shuswap

	Pattern One			Pattern Two			Pattern Three			Pattern Four	
	CQ	SB	Skt	PA	Ofo	Goj	BA	H	OG	Tzu	Shu
t^h⇔ʔ	✓			✓			✓	✓			
t'⇔ h	✓			✓			✓	✓		✓	
t^h⇔ɓ								✓			
ɗ⇔h								✓		✓	
h⇔ʔ	✓			✓			✓			✓	
t^h⇔h	*	*	*	✓	✓	✓	✓	✓			
$d^{ɦ}$⇔ h			*			✓					
t'⇔k^h	*			*			✓	✓	✓		
t^h⇔$g^{ɦ}$			*			*					
$d^{ɦ}$⇔$g^{ɦ}$			*			*					
t^h⇔k^h	*	*	*	*	*	*	✓	✓	✓		
t^h⇔$d^{ɦ}$			*			*					
ʔ⇔ɗ								*		✓	
ʔ⇔t'	*			*			*	*		✓	✓
t'⇔ɓ								*		✓	
ɓ⇔ɗ								*			
t'⇔k'	*			*			*	*	*	*	*
t'⇔t^h	*			*			*	*	*		
h⇔h	*	*	*	✓	✓		✓	✓		✓	
ʔ⇔ʔ								✓		✓	✓
t^h⇔t^h	*	*	*	✓	?	✓	✓	✓	✓		
$d^{ɦ}$⇔$d^{ɦ}$			*			✓					
t'⇔t'	*			✓			✓	✓	✓	✓	*
ɗ⇔ɗ								✓			

Table 3.5. Typology of laryngeal cooccurrence restrictions.

The typology of laryngeal cooccurrence restrictions is laid out in Table 3.5. This typology makes possible several significant observations. First, similarity is gradient across languages, and fits an implicational hierarchy. "Gradient" means that the point of

acceptable similarity (in terms of the graphics of Table 3.5, the discontinuity between checkmarks and asterisks in any column) is set differently in each pattern. The implicational hierarchy arises from the Pattern One restrictions being a subset of the Pattern Two restrictions, which are in turn a subset of the Pattern Three restrictions, and so on.

Second, some languages manifest the identity effect; this phenomenon takes one of two forms. First, if any pair of identical elements shown in the bottom half of Table 3.5 can cooccur, then it is likely that all such identical elements can cooccur. For example, if a language includes both aspirated and ejective segments in its inventory, then it will not allow a form such as t^{h}at^{h}a, while excluding t'at'a. This complete, or standard version of the identity effect is found in Peruvian Aymara, Gojri, Bolivian Aymara, Hausa, Old Georgian, and Tzutujil.

The identity effect can also take on a more complex form, as in Shuswap and (apparently) Ofo. In these languages, although two /h/s or two glottal stops can cooccur within morphemes, identical ejectives (and presumably identical aspirated stops) cannot.

Finally, the only case in which the subset effect does not hold concerns the cooccurrence of voiceless unaspirated stops with otherwise identical segments that differ in laryngeal features (t^{h}ati, t'ati, etc.). The restrictions on such forms are not shown in Table 3.5, but they are mentioned in the relevant sections on each language. In this respect, languages with the complete identity effect (i.e., most languages in Patterns Two,

Three, and Four) are actually stricter than languages without the complete identity effect (the languages of Pattern One, and Ofo and Shuswap).

The analysis presented in the following chapters accounts for all of these points.

CHAPTER 4: ANALYSIS – INTRODUCING CONSTRAINTS

In this chapter, I define and discuss the Optimality Theory constraints appealed to in the major portion of the analysis; in chapter 5, I discuss constraint rankings and illustrate important tableaux. Chapter 6 introduces more constraints and offers further analysis of the idiosyncratic facts of particular languages (for example, the leftness effect in Quechua and Aymara).

Optimality Theory (OT) is introduced and discussed in Prince and Smolensky 1993 and McCarthy and Prince 1993, 1994, and 1995. OT is a constraint-based theory of phonology in which the violability of constraints holds center stage. Constraints are held to be universal; differences among languages follow from two facts: (1) languages have distinct lexicons, and (2) constraints are differently ranked in different languages.

There are no derivations in OT; rather, a function Gen associates each input form to many candidate output forms. The selection of a winning output form is determined by constraint evaluation. The highest-ranking constraint is evaluated first; all output candidates violating this constraint are rejected. If more than one output candidate remains, then violations of the second highest-ranking constraint are considered. If no output candidate remains – as will be the case when all candidates violate the constraint under consideration – then in that case also the decision will devolve to the next highest-ranking constraint.

In OT, the essential tension within the grammar of a language is that between faithfulness constraints, requiring output forms to be similar to input forms, and phonological wellformedness constraints, which may require changes to input forms.

I use the Correspondence Theory approach to faithfulness (McCarthy and Prince 1995), rather than the earlier Containment Theory approach (McCarthy and Prince 1993). However, in place of the IDENT(feature) constraints of McCarthy and Prince 1995, I adopt the IDENTI→O(feature) and IDENTO→I(feature) constraints introduced in Pater (forthcoming). This means that faithfulness violations resulting from the insertion and deletion of features are calculated separately.

Fifteen of the constraints to which I appeal are shown in (1). In addition to these constraints, there are a variety of constraints that can be considered undominated, for the purposes of the material discussed here. These undiscussed constraints include the full range of IDENTI→O(entity) and IDENTO→I(entity) complexes (for example; IDENTI→O(sonorant), IDENTO→I(sonorant), etc.), and constraints that prevent sequential ordering relations from being changed.

Although laryngeal cooccurrence restrictions operate in some languages at the level of the word (as in Souletin Basque), the analysis provided here does not distinguish between those languages and languages in which the cooccurrence restrictions are restricted to the root or morpheme level. A theory of the domain of operation of cooccurrence restrictions remains a topic for future research.

(1)

1. **BEIDENTICAL – Segments should be identical.** One violation is assessed for every pair of non-identical segments. For simplicity's sake, only consonant violations are calculated in the tableaux provided here.

2-5. ***SIMILARITY – Consonants should not be similar.** One violation is assessed for every pair of consonants violating one of these constraints.

Each of the *SIMILARITY constraints refers to a different degree of similarity, and these constraints are subject to the universal ranking *SIM-4 >> *SIM-3 >> *SIM-2 >> *SIM-1. The point of acceptable similarity for each *SIMILARITY constraint is set as shown below. The languages listed are examples of languages for which that *SIMILARITY constraint is the active *SIMILARITY constraint – in other words, it is the lowest-ranked *SIMILARITY constraint that outranks (some of) the laryngeal correspondence constraints.

*SIM-1	*SIM-2	*SIM-3	*SIM-4
Cuzco Quechua	Peruvian Aymara	Bolivian Aymara	Tzutujil
identity	identity	identity	identity
t' ⇔ t^h	t' ⇔ t^h	t' ⇔ t^h	t' ⇔ t^h
t' ⇔ k'	t' ⇔ k'	t' ⇔ k'	t' ⇔ k'
ɓ ⇔ t'	ɓ ⇔ t'	ɓ ⇔ t'	
ʔ ⇔ t'	ʔ ⇔ t'	ʔ ⇔ t'	
ʔ ⇔ ɓ	ʔ ⇔ ɓ	ʔ ⇔ ɓ	
t^h ⇔ k^h	t^h ⇔ k^h		
t' ⇔ k^h	t' ⇔ k^h		
t^h ⇔ h			

6-8. **IDENTI→O(voice), IDENTI→O(spread glottis), IDENTI→O(constricted glottis)[1] – Laryngeal features present in the input should also be present on a corresponding segment (if any) in the output.** One violation is assessed for every laryngeal feature present in the input form that is not present in the output form.

9-11. **IDENTO→I(voice), IDENTO→I(spread glottis), IDENTO→I(constricted glottis) – Laryngeal features not present in the input should not be present on a corresponding segment (if any) in the output.** One violation is assessed for every laryngeal feature present in the output form that is not present in the input form.

12. **IDENTI→O(place) – Place features present in the input should also be present on a corresponding segment (if any) in the output.** One violation is assessed for every place feature present in the input form that is not present in the output form.

13. **IDENTO→I(place) – Place features not present in the input should not be present on a corresponding segment (if any) in the output.** One violation is assessed for every place feature present in the output form that is not present in the input form.

14. **MAX-IO(seg) – Segments present in the input should also be present in the output.** One violation is assessed for every segment present in the input form that is not present in the output form.

15. **DEP-IO(seg) – Segments present in the output should also be present in the input.[2]** One violation is assessed for every segment present in the output form that is not present in the input form.

---

[1] The IDENTI→O constraints listed in (6)-(8) and in (12) penalize feature deletion even when the host segment is not realized in the output form (i.e., is deleted). The IDENTO→I constraints listed in (9)-(11) and in (13) penalize feature insertion even when it results from epenthesis. As Pater points out, an alternative formulation of these constraints (which he ultimately adopts) would mandate featural faithfulness violations only when a corresponding segment is present. See Pater (forthcoming) for discussion.

[2] In Correspondence Theory, constraints labelled "MAX" penalize deletion; constraints labelled "DEP" penalize insertion.

In the following paragraphs I provide examples of how constraint violations are assigned, and suggest phonetic or processing bases for these constraints.

Regarding BEIDENTICAL, pata would garner one violation (because /p/ and /t/ are not identical), paska would incur three violations (because none of the pairs {/p/, /s/}; {/s/, /k/}, or {/p/, /k/} are identical), p'asp'a would earn two violations (neither of the two {/p'/, /s/} pairs are identical, but the {/p'/, /p'/} pair *is* identical), and t'at'a would involve no BEIDENTICAL violations (/t'/ and /t'/ are identical).

The reader may be wondering about input forms such as pata. What would prevent this form from being realized as papa or tata, thus avoiding a BEIDENTICAL violation? These forms will be ruled out if either IDENTI→O(place) or IDENTO→I(place) outranks BEIDENTICAL.

The BEIDENTICAL constraint is likely to have an articulatory basis: programming two identical segments is probably easier than programming two different segments. Linguistic phenomena such as morphological reduplication, segment harmony processes in child speech, and reduplication processes among speakers with language deficits may offer other illustrations of this constraint in action.

The *SIMILARITY constraints penalize segments for being similar. Note that these constraints, as I have written them, only refer to the similarities of segments concerned in laryngeal cooccurrence restrictions. This is a simplification. It is probable that these constraints actually refer to a single quantity, auditory similarity. In this case, all similar segments would belong in the hierarchy. It would include, for example, pairs of segments

sharing place of articulation. If this is the case, then one could write a more complete set of *SIMILARITY constraints. Whether a particular language had, for example, place or laryngeal cooccurrence restrictions would be determined by the relative rankings of the various correspondence constraints (i.e. IDENTI→O(place), IDENTO→I(place), IDENTI→O(voice), IDENTO→I(voice), etc.) and the *SIMILARITY constraints.[3]

*SIMILARITY penalizes the cooccurrence of two or more similar segments. Notice that identical segments, being maximally similar, also count as *SIMILARITY violations. The identity pole of the *SIMILARITY hierarchy works in a direction opposite to the constraint BEIDENTICAL: identical segments are disfavored because they are maximally similar – they will always involve *SIMILARITY violations. In languages with highly-ranked BEIDENTICAL, these segments will cooccur anyway.

Like BEIDENTICAL violations, *SIMILARITY violations are counted in a pairwise fashion. The output form p'asp'a produces a *SIMILARITY violation regardless of which *SIMILARITY constraint is active, due to the presence of identical bilabial ejectives within the morpheme. The form pata would not be assessed any *SIMILARITY violations (according to the formulation given here), regardless of which *SIMILARITY constraint is active. The form p'atʰa would involve a *SIMILARITY violation if either *SIM-1 or *SIM-2 were active. Sample violations of the four *SIMILARITY constraints are given in (2).

[3] In earlier work (MacEachern 1996, 1997), I developed *SIMILARITY as a scalar constraint, describing seven degrees of similarity based on auditory features. In MacEachern 1997, I also noted that it would also be possible to formalize *SIMILARITY as a set of constraints conjoined in the manner of Flemming 1995.

(2)

	*SIM-1	*SIM-2	*SIM-3	*SIM-4
pata				
hat^{h}a	*			
p'at^{h}a	*	*		
p^{h}at^{h}a	*	*		
ʔat'a	*	*	*	
p'at'a	*	*	*	*
t^{h}at'a	*	*	*	*
p^{h}asp^{h}a	*	*	*	*
p'asp'a	*	*	*	*
papa	*	*	*	*

I suggest that the points on the similarity scale are universally ranked with respect to one another, and that this ranking follows from auditory similarity. The crosslinguistic typology of laryngeal cooccurrence restrictions is constrained by the similarity scale, and this accounts for the implicational hierarchy manifested in Table 3.5. Unfortunately, I do not have experimental evidence supporting the relative auditory similarities of the elements in question; my thoughts on this point are presented in the next few paragraphs.

In chapter 1, I offered the similarity continuum shown in (3), and pointed out that in Cuzco Quechua, elements that are adjacent in (3) do not cooccur within roots.

(3)

h------Ch------C'------ʔ

Furthermore, none of the elements shown in (3) coexist with identical copies of themselves (*halha, *t'ult'u, etc.). However, elements that are separated by at least one other element *do* cooccur: for example, hutk'u 'tube-like hole', ʔantʃ^{h}a 'very'.

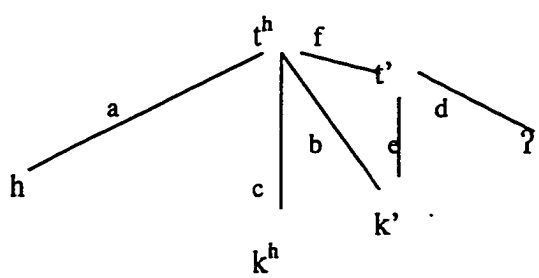
In (3), the four elements under discussion are shown with equal intervals intervening. Of course, these elements are not necessarily equal distances apart on the scale of perceptual similarity. Suppose that we found through the construction of audition-based confusion matrices that the actual similarity distances of the elements in (3) were something like what is shown in (4).

(4)

h----------Ch----C'---------ʔ

A scale like the above seems plausible: aspirated and ejective stops do sound more alike than /h/ and aspirated stops. Such a diagram would also account for why /h/ and aspirated stops can cooccur in the languages of Patterns Two and higher, while aspirated stops and ejective stops, for example, cannot: /h/ and Ch are actually less similar than Ch and C'. Now consider what might happen if we add in the effects of place of articulation.

(5)



a > b > c > d > e > f

The graphic in (5) illustrates the significance of place of articulation, which was not represented in (4). Some of the relative distances represented in (5) are uncontroversial: certainly two aspirated stops (separated by distance c) are more similar than are /h/ and an aspirated stop (separated by distance a), for instance. Certainly identical elements (found in the five rows at the bottom of the hierarchy in Table 3.5), which would be represented by a distance of zero in (5), are more similar than non-identical elements.

In other respects, the similarity gradient is less obvious: heterorganic aspirated stops (c) might be less similar to each other than are heterorganic ejective stops (e), for example. Formant transition information is available in the aspirated transitions of **t^h** ⇔ **k^h** forms; such information is not present in the silent transitions of **t'** ⇔ **k'** forms. However, it might be the case that the high-amplitude release bursts of ejective consonants are very informative, perhaps even more so than the release bursts and aspirated transitions of aspirated stops. If this were the case, then the relative locations of **t^h** ⇔ **k^h** and **t'** ⇔ **k'** in Table 3.5 – and the relative distances c and e in (5) – would not support an understanding of this hierarchy that is based on auditory similarity. Clearly, experimental evidence is needed to confirm the relative similarities of the elements in the hierarchy.[4]

---

[4] Other investigators have proposed similarity metrics that refer to phonological features. Recall that both Cuzco Quechua and Peruvian Aymara prohibit aspirated stops from cooccurring with ejective stops, but allow voiceless unaspirated stops to cooccur with both aspirated and ejective stops. If a phonological feature-based similarity metric is proposed to account for this data, it would at least have to refer to some feature that covers both aspiration and ejectiveness on stops – perhaps voice onset time.

Moving on to the other constraints listed in (1), constraints 6-11 (IDENTI→O(voice), IDENTI→O(spread glottis), IDENTI→O(constricted glottis), IDENTO→I(voice), IDENTO→I(spread glottis), and IDENTO→I(constricted glottis)) are Correspondence Theory constraints that enforce featural identity between the input and output forms.[5] The IDENTI→O(feature) and IDENTO→I(feature) constraint complexes penalize the deletion and insertion, respectively, of laryngeal features. One violation is incurred for each omitted or inserted feature specification. For example, the input-output pair /sapa/-sap'a would incur one IDENTO→I(constricted glottis) violation; the input-output pair /sap'a/-sapa would incur one IDENTI→O(constricted glottis) violation.

IDENTI→O(place) and IDENTO→I(place) work analogously to the laryngeal correspondence constraints described above. The input-output pair /t'ap'a/-t'at'a would involve one IDENTI→O(place) violation and one IDENTO→I(place) violation. The segments /h/ and /ʔ/ are understood to be placeless, so the input-output pair /tapa/-hapa, for example, would involve one IDENTI→O(place) violation and no IDENTO→I(place) violations.

MAX-IO(seg) penalizes the deletion of segments; DEP-IO(seg) penalizes epenthesis. The input-output pair /salp'a/-sap'a would involve one MAX-IO(seg) violation; the pair /sap'a/-salp'a would involve one DEP-IO(seg) violation.

---

[5] In MacEachern 1996, 1997, I used a single laryngeal faithfulness constraint, IDENT(lar). This constraint required corresponding segments in input and output forms to be identical with respect to laryngeal features. It collapsed distinctions between deletion and insertion effects, and between laryngeal features representing voicing, aspiration, and ejectiveness.

In the following chapter, I discuss the rankings of these constraints within each of the languages under consideration, and provide explanatory tableaux.

CHAPTER 5: ANALYSIS – TABLEAUX AND RANKINGS

In the analysis presented here, the different cooccurrence restrictions in the languages under discussion mainly follow from two conditions: first, whether or not BEIDENTICAL outranks all of the *SIMILARITY constraints, and second, which *SIMILARITY constraint (*SIM-4, *SIM-3, *SIM-2, or *SIM-1) is active. By "active", I mean the constraint that determines the point of unacceptable similarity in the language under investigation. This will be the strictest *SIMILARITY constraint (*SIM-1 is stricter than *SIM-2, which is in turn stricter than *SIM-3, etc.) that outranks any of the laryngeal Correspondence constraints (i.e. IDENTI→O(spread glottis), IDENTO→I(spread glottis), etc.).

In the following sections, I provide one possible constraint ranking for each language, and offer tableaux illustrating important points of the analysis. Other rankings may also generate the cooccurrence restrictions under discussion, and in any case I am not providing tableaux in order to motivate the rankings suggested here. Rather, I am providing tableaux in order to illustrate how the constraint ranking of each language generates the cooccurrence restrictions of the grammar. In order to do this, I assume various hypothetical input forms and show that output candidates that violate the cooccurrence restrictions are not selected as winning candidates: given the constraint rankings provided here, all winning candidates will obey the cooccurrence restrictions of the language in question.

All constraint rankings were checked with software written by Bruce Hayes. This software allows the user to enter input forms, output candidates, and constraint violations, and to select a winning output form. If a constraint ranking consistent with these parameters is available, the software will provide it. All constraint rankings given here produced the correct results for each language, but the results of course depend on my having chosen the appropriate sets of input and output forms. My methods for selecting these sets are described in the following paragraphs.

Consider Cuzco Quechua. In this language, there are five types of relevant segments: /ʔ/, /h/, aspirated stops, ejective stops, and voiceless unaspirated stops. We examine forms that include pairs of segments of the relevant types, so there are 15 possible combinations of these elements within and across classes (I am abstracting away from vowels here): /ʔ, ʔ/, /h, h/, /t^h, t^h/, /t', t'/, /t, t/, /ʔ, h/, /ʔ, t^h/, /ʔ, t'/, /ʔ, t/, /h, t^h/, /h, t'/, /h, t/, /t^h, t'/, /t^h, t/, /t', t/ (order is not relevant here). Place of articulation becomes relevant when elements with supralaryngeal Place features are combined. In Cuzco Quechua, this adds six forms to the list of relevant input forms: /t^h, k^h/, /t', k'/, /t, k/, /t^h, k'/, /t^h, k/, /t', k/. (Note that *which* places of articulation are used is irrelevant; what matters is whether segments are heterorganic or homorganic.)

In Cuzco Quechua, we can ignore the one form that includes two glottal stops, because glottal stop can only appear in initial position in this language. We are left with 20 relevant input forms.

Souletin Basque has /h/, voiceless aspirated stops, voiced stops, and voiceless unaspirated stops. There are 16 relevant combinations of these elements: /h, h/, /h, t^h/, /h, d/, /h, t/, /t^h, t^h/, /t^h, d/, /t^h, t/, /d, d/, /d, t/, /t, t/, /t^h, k^h/, /t^h, g/, /t^h, k/, /d, g/, /d, k/, /t, k/.

Similar calculations for the other languages determine the sets of input candidates given below.

Sanskrit: /h, h/, /h, t^h/, /h, d/, /h, $d^ɦ$/, /h, t/, /t^h, t^h/, /t^h, d/, /t^h, $d^ɦ$/, /t^h, t/, /d, d/, /d, $d^ɦ$/, /d, t/, /$d^ɦ$, $d^ɦ$/, /$d^ɦ$, t/, /t, t/, /t^h, k^h/, /t^h, g/, /t^h, $g^ɦ$/, /t^h, k/, /d, g/, /d, $g^ɦ$/, /d, k/, /$d^ɦ$, $g^ɦ$/, /$d^ɦ$, k/, /t, k/.

Peruvian Aymara: same as Cuzco Quechua.

Ofo: /h, h/, /t^h, h/, /t^h, t^h/, /t^h, k^h/, /t, h/, /t^h, t/, /t^h, k/, /t, k/, /t, t/.

Gojri: /h, t^h/, /h, d/, /h, t/, /t^h, t^h/, /t^h, d/, /t^h, t/, /d, d/, /d, t/, /t, t/, /t, k/, /t, k^h/, /t, g/, /t^h, k^h/, /t^h, g/, /d, g/ (/h, h/ is not possible because /h/ only occurs in initial position).

Bolivian Aymara: same as Cuzco Quechua and Peruvian Aymara.

Hausa: /t', t^h/, /ʔ, t'/, /ʔ, t^h/, /ʔ, ɓ/, /ʔ, b/, /h, t'/, /h, t^h/, /h, ɓ/, /h, b/, /t', ɗ/, /t', d/, /t^h, ɗ/, /t^h, d/, /ɓ, b/, /t', t'/, /t^h, t^h/, /ɓ, ɓ/, /b, b/, /t', k^h/, /t', ɓ/, /t', b/, /t^h, ɓ/, /t^h, b/, /ɓ, d/, /t', k'/, /t^h, k^h/, /ɓ, ɗ/, /b, d/ (/ʔ/ and /h/ don't cooccur medially in native morphemes, so the three possible forms /ʔ, ʔ/, /h, h/, and /ʔ, h/ are ruled out).

Shuswap: /ʔ, ʔ/, /ʔ, t'/, /t', t'/, /t', k'/, /ʔ, t/, /t', t/, /t', k/, /t, k/, /t, t/.

Old Georgian: /t', t'/, /t', k'/, /t^h, k^h/, /t^h, t^h/, /b, b/, /b, d/, /k^h, k'/, /t^h, k'/, /k', g/, /ts', b/, /g, k^h/, /g, t^h/.

Tzutujil: /ʔ, h/, /ʔ, t'/, /ʔ, ʈ/, /ʔ, ɓ/, /h, t'/, /h, ʈ/, /h, ɓ/, /t', ʈ/, /t', ɗ/, /t, ɗ/, /ʔ, ʔ/, /h, h/, /t', t'/, /t, ʈ/, /ɓ, ɓ/, /t', k'/, /t, k/, /ɓ, ɗ/, /t', k/, /t', ɓ/, /t, ɓ/.

Output candidate sets were generated by changing laryngeal and/or place features of one or both of the input segments, and/or by deleting one of these segments. Here I will just provide one example of how output candidate sets were built, using the Peruvian Aymara input form /t', k'/.

The input segments are ejectives, so there are three relevant alternations: we could subtract a [constricted glottis] feature (making the segments plain voiceless unaspirated), we could subtract a [constricted glottis] feature and add a [spread glottis] feature (making the segments voiceless aspirated), or we could change the place of articulation of one of the segments in order to render them homorganic. Combining these alternations with the possibility of leaving the segments unaltered produces 18 possible output candidates: {t', k'}, {t', k}, {t', k^h}, {t, k'}, {t, k}, {t, k^h}, {t^h, k'}, {t^h, k}, {t^h, k^h}, {t', t'}, {t', t}, {t', t^h}, {t, t'}, {t, t}, {t, t^h}, {t^h, t}, {t^h, t^h}, {t^h, t'}. We could also delete one or both segments of the input; nothing would be gained by altering one segment when the other is deleted, and IDENTI→O(feature) violations would be incurred, so the only relevant inputs involving MAX-IO(seg) violations would be as follows: {Ø, k'}, {t', Ø}, {Ø, Ø}. Finally, we could replace one or both of the segments with /h/ or with /ʔ/: {h, k'}, {ʔ, k'}, {h, k}, {ʔ, k}, {h, k^h}, {ʔ, k^h}, {h, Ø}, {ʔ, Ø}, {h, h}, {h, ʔ}, {ʔ, ʔ}.

The two candidates {t', k} and {t, k'} are equivalent, given the set of constraints discussed so far, so we will remove {t, k'} from consideration (any ranking would assign it the same number of constraint violations as {t', k}, the actual winner; {t', k} wins due to the leftness effect, which is covered in chapter 6). The six candidates {t, k^h}, {t^h, k'}, {Ø, k'}, {t, t'}, {t^h, t'}, and {t, t^h} can be dropped for similar reasons.

We are left with the following 25 output candidates: {t', k'}, {t', k}, {t', k^h}, {t, k}, {t^h, k}, {t^h, k^h}, {t', t'}, {t', t}, {t', t^h}, {t, t}, {t^h, t}, {t^h, t^h}, {t', Ø}, {Ø, Ø}, {h, k'}, {ʔ, k'}, {h, k}, {ʔ, k}, {h, k^h}, {ʔ, k^h}, {h, Ø }, {ʔ, Ø }, {h, h}, {h, ʔ}, {ʔ, ʔ}.

Selection of the winning output candidate is simple in cases where the input candidate does not violate the cooccurrence restrictions of the language, as in Peruvian Aymara /h t'/, /t k/, and /h t^h/. In each case, the grammar of the language will select the faithful candidate as the winning output form. When the input form does violate the cooccurrence restrictions of the language, as in Peruvian Aymara /t' k'/, some output candidate other than the faithful candidate (in this case, {t', k'}) must be the winning output. The optimal candidate must obey the cooccurrence restrictions of the language, and it must be chosen so that a single constraint ranking can be determined for the language. In the case of /t' k'/, I have chosen to make {t', k} the winning output. When the constraints are ranked, this will result in the demotion of IDENTI→O(constricted glottis) to a position below the active *SIM-2.

In order to help the reader follow the analysis, an overview of the data patterns (Table 3.5) is reproduced below as Table 5.1.

	Pattern One			Pattern Two			Pattern Three			Pattern Four	
	CQ	SB	Skt	PA	Ofo	Goj	BA	H	OG	Tzu	Shu
$t^h \Leftrightarrow ʔ$	✓			✓			✓	✓			
$t' \Leftrightarrow h$	✓			✓			✓	✓		✓	
$t^h \Leftrightarrow ɓ$								✓			
$ɗ \Leftrightarrow h$								✓		✓	
$ʰ \Leftrightarrow ʔ$	✓			✓			✓			✓	
$t^h \Leftrightarrow h$	*	*	*	✓	✓	✓	✓	✓			
$d^ɦ \Leftrightarrow h$			*			✓					
$t' \Leftrightarrow k^h$	*			*			✓	✓	✓		
$t^h \Leftrightarrow g^ɦ$			*			*					
$d^ɦ \Leftrightarrow g^ɦ$			*			*					
$t^h \Leftrightarrow k^h$	*	*	*	*	*	*	✓	✓	✓		
$t^h \Leftrightarrow d^ɦ$			*			*					
$ʔ \Leftrightarrow ɗ$								*		✓	
$ʔ \Leftrightarrow t'$	*			*			*	*		✓	✓
$t' \Leftrightarrow ɓ$								*		✓	
$ɓ \Leftrightarrow ɗ$								*			
$t' \Leftrightarrow k'$	*			*			*	*	*	*	*
$t' \Leftrightarrow t^h$	*			*			*	*	*		
$h \Leftrightarrow h$	*	*	*	✓	✓		✓	✓		✓	
$ʔ \Leftrightarrow ʔ$								✓		✓	✓
$t^h \Leftrightarrow t^h$	*	*	*	✓	?	✓	✓	✓	✓		
$d^ɦ \Leftrightarrow d^ɦ$			*			✓					
$t' \Leftrightarrow t'$	*			✓			✓	✓	✓	✓	*
$ɗ \Leftrightarrow ɗ$								✓			

Table 5.1. Typology of laryngeal cooccurrence restrictions.

5.1 Pattern One

5.1.1 Cuzco Quechua

First, consider Cuzco Quechua. This language has voiceless unaspirated, voiceless aspirated, and ejective segments, /h/, and glottal stop (which occurs only in initial position). Data from this language is discussed in 3.1.1; the language includes the following cooccurrence restrictions: there is only one ejective or aspirated stop per morpheme, /h/ does not cooccur with /h/ or with aspirated stops, and glottal stop does not cooccur with ejectives.

One constraint ranking that will generate these restrictions is shown in (1) (I include IDENTI→O(voice) and IDENTO→I(voice) in the ranking given below, but note that voiced obstruents do not exist in Cuzco Quechua).

(1)

MAX-IO(seg) DEP-IO(seg) IDENTI→O(place) IDENTI→O(voice) IDENTO→I(voice) IDENTO→I(spread glottis) IDENTO→I(constricted glottis)

↓

*SIM-4 *SIM-3 *SIM-2 *SIM-1

↓

BEIDENTICAL IDENTI→O(spread glottis) IDENTO→I(place) IDENTI→O(constricted glottis)

For ease of presentation, the unviolated MAX-IO(seg), DEP-IO(seg), IDENTI→O(place), IDENTI→O(voice), IDENTO→I(voice), and IDENTO→I(constricted glottis) constraints are omitted from the following tableaux. Since no other constraints are interleaved between the *SIMILARITY constraints, only the active constraint *SIM-1 is shown; additional violations due to the presence of the other *SIMILARITY constraints are irrelevant to the selection of winning candidates.

The tableau in (2) assumes an input form with two ejective stops. The constraint ranking given will ensure that the output candidate that most transparently matches this input – *t'ak'a – cannot surface, because it involves a *SIMILARITY violation. The winning candidate, t'aka, tolerates an IDENTI→O(constricted glottis) violation but avoids a *SIMILARITY violation.

(2)

	/t'ak'a/	IDO→I(sg)	*SIM-1	BEIDENTICAL	IDI→O(sg)	IDO→I(place)	IDI→O(cg)
☞	t'aka						*
	t'ak'a		*!				

In Cuzco Quechua, the leftness effect (discussed in chapter 6) establishes that it is t'aka rather than tak'a that is the winning output form. Aspirated stops work similarly to ejective segments. The constraint ranking proposed here will produce the output t^haka for the input /t^hak^ha/.

Another restriction of Cuzco Quechua is that which prohibits aspirated stops from cooccurring with ejective stops. A tableau establishing this is shown in (3).

(3)

	/k^nut'u/	IDO→I(sg)	*SIM-1	BEIDENTICAL	IDI→O(sg)	IDO→I(place)	IDI→O(cg)
☞	k^nutu			*			*
	k^nut'u		*!	*			
	kutu			*	*!		*

As (3) shows, one of the laryngeal features will be sacrificed in order to avoid a *SIMILARITY violation.

The tableau in (4) shows how forms with two /h/s might be ruled out.

(4)

	/haha/	IDO→I(sg)	*SIM-1	BEIDENTICAL	IDI→O(sg)	IDO→I(place)	IDI→O(cg)
☞	hata						
	haha		*!				

The haha output involves a *SIMILARITY violation. The hata output avoids a *SIMILARITY violation but incurs both IDENTI→O(spread glottis) and IDENTO→I(place)

violations, because an aspiration feature is deleted and a place feature is inserted. This ranking is one way of accounting for the absence of forms containing two /h/s.

The segment /h/ is also prohibited from cooccurring with aspirated stops, as shown in (5). In this tableau, I include IDENTO→I(constricted glottis); IDENTO→I(spread glottis) is omitted due to space limitations.

(5)

	/hat^{h}un/	IDO→I(cg)	*SIM-1	BEIDENTICAL	IDI→O(sg)	IDO→I(place)	IDI→O(cg)
☞	hatun						
	hat^{h}un		*!				
	ʔat^{h}un	*!					

Again, one of the laryngeal features must be sacrificed in order to satisfy *SIMILARITY. The aspiration is dropped from /t^h/, resulting in a low-ranked IDENTI→O(spread glottis) violation.[1]

Glottal stop may not cooccur with ejectives, as shown in (6).

(6)

	/ʔak'a/	IDO→I(sg)	*SIM-1	BEIDENTICAL	IDI→O(sg)	IDO→I(place)	IDI→O(cg)
☞	ʔaka						
	ʔak'a		*!				

Tableau (6) shows what should be by now a familiar process: a laryngeal feature is omitted in order to avoid a *SIMILARITY violation.

[1] Interestingly, given an input form such as /t^{h}ahun/, the winning output will be t^{h}atun rather than *tahun. This input-output pair requires reference to constraints governing the leftness effect, which is discussed in chapter 6.

The tableau in (7) shows that Cuzco Quechua ejectives *are* allowed to cooccur with /h/.

(7)

	/hitʃ'i/	IDO→I(sg)	*SIM-1	BEIDENTICAL	IDI→O(sg)	IDO→I(place)	IDI→O(cg)
☞	hitʃ'i			*			
	hitʃi			*			*!

The winning candidate involves only a BEIDENTICAL violation, since /h/ and /tʃ'/ are of course not identical. The form hihi is not listed in the tableau, but it would involve violations of the undominated constraints IDENTI→O(place) and IDENTO→I(spread glottis); the form tʃ'itʃ'i would involve a violation of undominated IDENTO → I(constricted glottis).

Finally, we must consider the set of forms in Cuzco Quechua that *do* have identical stops – forms such as tuta 'dark' and kunka 'neck'. The tableau in (8) shows how these forms are allowed to cooccur, while forms such as $t^h ut^h a$ and k'unk'a are ruled out; the input form /papa/ is used an an example. In this tableau, I include the undominated constraint IDENTI→O(place), and omit irrelevant IDENTI→O(constricted glottis).

(8)

	/papa/	IDI→O(place)	IDO→I(sg)	*SIM-1	BEIDENTICAL	IDI→O(sg)	IDO→I(place)
☞	papa						
	p^napa		*!				
	pata	*!					*

This completes my review of the basic analysis of the Cuzco Quechua data. Additional analysis of the leftness effect is provided in 6.1.

5.1.2 Souletin Basque

The data of Souletin Basque were discussed in 3.1.2. As the reader may recall, this language differs from Cuzco Quechua in three relevant respects. First, ejective and glottal stops are not present in the segment inventory of the language. Second, voiced stops *are* present. The language has three stop series: voiceless unaspirated, voiceless aspirated, and voiced. Finally, the morpheme-level cooccurrence restrictions of this Basque dialect are mirrored by productive processes that apply across morpheme boundaries. I do not discuss those processes here.

As in Cuzco Quechua, aspirated stops may not occur with one another or with /h/, and two /h/s do not cooccur. Forms with two voiceless unaspirated stops (for example, kiskili 'blister') are allowed. The voiced stops of Souletin Basque, however, introduce a new dimension into the analysis. Forms with a voiced stop and another stop differing in manner – for example, /bapa/ and /pʰaba/ – are allowed to occur, as are forms with two voiced stops (for example, dendari 'dressmaker').

The ranking given in (1) for Cuzco Quechua will also serve for Souletin Basque, although IDENTI→O(constricted glottis) can be elevated in the hierarchy, since ejective segments and glottal stop do not exist in Basque. This ranking is reproduced in (9).

(9)

MAX-IO(seg)
DEP-IO(seg)
IDENTI→O(place)
IDENTI→O(voice)
IDENTO→I(voice)
IDENTO→I(spread glottis)
IDENTI→O(constricted glottis)
IDENTO→I(constricted glottis)

↓

*SIM-4
*SIM-3
*SIM-2
*SIM-1

↓

BEIDENTICAL
IDENTI→O(spread glottis)
IDENTO→I(place)

In this section I will present just one Basque tableau. The tableau in (10) shows how identical voiced stops can cooccur in this dialect. Again, only a subset of the constraints are shown.

(10)

	/dada/	IDI→O(place)	IDI→O(voice)	IDO→I(voice)	*SIM-1	BEIDENTICAL	IDO→I(place)
☞	dada						
	tada		*!				
	daba	*!					*

The output with identical voiced stops is selected as the winning candidate, despite involving a *SIMILARITY violation.

Souletin Basque is discussed further in 6.1.

5.1.3 Sanskrit

The Sanskrit data are discussed in 3.1.3. The language has four stop series: voiceless unaspirated, voiceless aspirated, voiced unaspirated, and voiced aspirated. The laryngeal fricative also exists. No aspirated segments – whether /h/, voiceless aspirated stops, or voiced aspirated stops – can cooccur within morphemes.

The existence of a fourth stop series in Sanskrit (namely voiced aspirated stops) allows for several input possibilities beyond those discussed for Souletin Basque. However, the ranking given for that language will also generate the cooccurrence restrictions of Sanksrit. (As in Souletin Basque, of course, the absence of glottalic segments in the language means that IDENTI→O(constricted glottis) is not conclusively dominated, as it is in Cuzco Quechua.) I do not offer any tableaux here.

5.2 Pattern Two

5.2.1 Peruvian Aymara

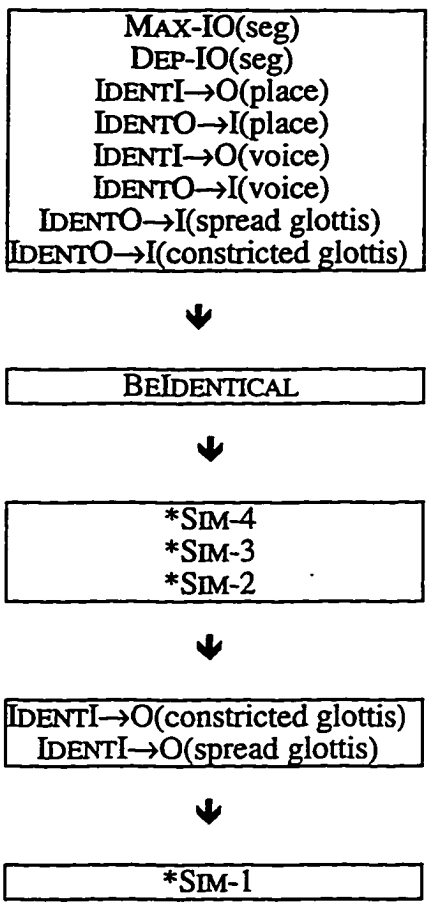
Peruvian Aymara has virtually the same segment inventory as Cuzco Quechua: the language has /h/, glottal stop (which is noncontrastive and occurs only in initial position), and three stop series: voiceless unaspirated, voiceless aspirated, and ejective.

The data of Peruvian Aymara is discussed in 3.2.1. In this language, unlike Cuzco Quechua, aspirated stops can cooccur with identical aspirated stops, and ejectives can cooccur with identical ejectives. Two /h/s can also cooccur within morphemes, and aspirated stops can cooccur with /h/. Finally, morphemes with homorganic aspirated and

voiceless unaspirated stops, and morphemes with homorganic ejective and voiceless unaspirated stops, are rare. These are all the respects in which Peruvian Aymara differs from Cuzco Quechua. The cooccurrence restrictions of Peruvian Aymara prohibit the cooccurrence within a morpheme of heterorganic ejectives, heterorganic aspirated stops, and ejectives and aspirated stops. In addition, glottal stop cannot cooccur with ejectives.

The identity effect (all identical segments may cooccur) is produced when BEIDENTICAL outranks the active *SIMILARITY constraint. One ranking that produces the data patterns summarized above is shown in (11). Undominated and other irrelevant constraints are omitted from the tableaux that follow.

(11)



The tableau in (12) shows how heterorganic ejectives are prohibited in Peruvian Aymara. In this tableau, a hypothetical input form containing a pair of heterorganic ejectives is mapped onto an output form containing only one ejective.

(12)

	/q'at'a/	BEIDENTICAL	*SIM-2	IDI→O(cg)	IDI→O(sg)	*SIM-1
☞	q'ata	*				
	q'at'a	*	*!			*

Both of the candidates shown in (12) involve BEIDENTICAL violations, but q'ata avoids a *SIMILARITY violation, so it emerges the winner. The alternative candidates

t'at'a and q'aq'a (not shown) would involve higher-ranked IDENTI→O(place) and IDENTO→I(place) violations.

The tableau in (13) shows how in Peruvian Aymara, unlike in Cuzco Quechua, identical ejectives may cooccur.

(13)

	/p'ap'i/	BEIDENTICAL	*SIM-2	IDI→O(cg)	IDI→O(sg)	*SIM-1
☞	p'ap'i		*			*
	p'api	*!		*		
	papi		*	**!		*

The output form with two identical ejectives, p'ap'i, satisfies BEIDENTICAL. It violates *SIMILARITY, but incurs no IDENTI→O(constricted glottis) violations, so it emerges the winner. Tableaux analogous to those in (12) and (13) could be produced to show that aspirated stops work the same way.

The tableau in (14) illustrates how forms with two /h/s may occur, another respect in which Peruvian Aymara differs from the Pattern One languages. The candidate hata, of course, would also violate IDENTI→O(place) (not shown).

(14)

	/haha/	BEIDENTICAL	*SIM-2	IDI→O(cg)	IDI→O(sg)	*SIM-1
☞	haha					
	hata	*!				

The tableau in (15) illustrates yet another way in which Peruvian Aymara differs from Cuzco Quechua. Aspirated stops and /h/ can cooccur, because the cooccurrence of

these elements does not violate *SIM-2, which is the active *SIMILARITY constraint in Peruvian Aymara.

(15)

	/hat^{h}i/	BEIDENTICAL	*SIM-2	IDI→O(cg)	IDI→O(sg)	*SIM-1
☞	hat^{h}i	*				*
	?at^{h}i	*			*!	
	hati	*			*!	

The successful candidate hat^{h}i involves only a BEIDENTICAL violation and a low-ranked *SIM-1 violation.

One final respect in which Peruvian Aymara differs from Cuzco Quechua is in the absence of forms such as k'alka and t^{h}ata. This restriction appears in both Peruvian and Bolivian Aymara, and it represents the only case in which the restrictions of these languages are not a subset of the tighter set of restrictions found in Cuzco Quechua. The tableau in (16) suggests that the effect arises because a laryngeal feature will be sacrificed in order to achieve consonantal identity. In other words, forms such as t^{h}ati and t'ata will be absent in just those languages that have the (complete) identity effect.[2] This does not happen in Cuzco Quechua, where the laryngeal Correspondence constraints are not outranked by BEIDENTICAL.

[2] Such forms are predicted to be present in languages that have the partial, or incomplete version of the identity effect. See 5.2.2 (Ofo) and 5.4.1 (Shuswap) for more information.

(16)

	/t^{n}ati/	BEIDENTICAL	*SIM-2	IDI→O(cg)	IDI→O(sg)	*SIM-1
	t^{n}ati	*!				
	tati		*		*!	
☞	t^{n}at^{n}i		*			*

Consideration of underlying forms such as /t'ati/, of course, would show that ejectives and homorganic voiceless unaspirated stops work the same way.

Analysis of the leftness effect in Peruvian Aymara is given in 6.1.

5.2.2 Ofo

The Ofo data were discussed in 3.2.2. The language has two simple voiceless series, unaspirated and aspirated. As in Peruvian Aymara, non-identical aspirated obstruents are not permitted to cooccur within words, but /h/ is allowed to cooccur with aspirated obstruents. However, the language differs from Peruvian Aymara in two respects: first, although forms with two /h/s are found, forms with identical aspirated stops are not attested. Recall, however, that I argued in 3.2.2.4 that given the limited amount of data we have on this language, we do not know whether or not the lack of morphemes containing identical aspirated stops is accidental. Second, aspirated segments do cooccur with their unaspirated counterparts (for example, po'p^{h}uti 'to swell or puff out').

Thus, there are two facts of which we are certain: Ofo had forms containing two /h/s, and Ofo had forms containing homorganic voiceless aspirated and voiceless unaspirated obstruents. There is one point on which we are uncertain: namely, whether

or not the Ofo language had forms containing identical aspirated obstruents (for example, t^{h}at^{h}i).

The analysis given here indicates that Ofo cannot have had such forms. The only constraint rankings that will allow forms with two /h/s *and* forms of type pop^{h}uti will make forms with identical aspirated obstruents impossible. I conclude that Ofo must have been a language characterized by the partial, or incomplete version of the identity effect.

In order to show why forms such as ahihi, t^{h}at^{h}i, and pop^{h}uti are incompatible, let me assume for the moment that Ofo *did* allow forms with identical aspirated obstruents, and that the lack of such forms in Dorsey and Swanton 1912 is accidental.

First, consider the tableau in (17).

(17)

	/t^{h}ahi/	BEIDENTICAL	*SIM-2	IDI→O(sg)	*SIM-1
☞	t^{h}ahi	*			*
	tahi	*		*!	

Ofo allows aspirated obstruents to cooccur with /h/, so t^{h}ahi must be selected as the winner. This tableau shows that IDENTI→O(spread glottis) must outrank *SIM-1, if the correct output (t^{h}ahi) is to be selected.

Now consider an input form with non-identical aspirated obstruents. Such forms do not exist in Ofo, so the tableau in (18) shows that *SIM-2 (and, by extension, all *SIMILARITY constraints that are less strict) must outrank IDENTI→O(spread glottis).

(18)

	/t^nak^ni/	BEIDENTICAL	*SIM-2	IDI→O(sg)	*SIM-1
	t^nak^ni	*	*!		
☞	t^naki	*			

At this point, we have established the ranking *SIM-2 >> IDENTI→O(spread glottis) >> *SIM-1. The tableau in (19) shows that if an input with identical aspirated obstruents is to emerge with its laryngeal features intact, then BEIDENTICAL must outrank *SIM-2.

(19)

	/t^nat^ni/	BEIDENTICAL	*SIM-2	IDI→O(sg)	*SIM-1
☞	t^nat^ni				
	t^nati	*!			

The ranking proposed for handling this data set (which assumes the presence in Ofo of forms containing identical aspirated obstruents) is now BEIDENTICAL >> *SIM-2 >> IDENTI→O(spread glottis) >> *SIM-1. Now consider the tableau in (20), illustrating two possible output candidates for the input form /t^hati/. The exclamation point preceding the selection finger indicates that the tableau generates the wrong output candidate.

(20)

	/t^nati/	BEIDENTICAL	*SIM-2	IDI→O(sg)	*SIM-1
	t^nati	*!			
!☞	tati				

This constraint ranking will wrongly select <u>tati</u> (or, alternatively, <u>t^hat^hi</u>, depending on the location of IDENTO→I(spread glottis) within the constraint hierarchy) as the

winning output candidate. There is no way to derive t^{h}ati from /t^{h}ati/, and more generally, no way to produce forms containing homorganic aspirated and voiceless unaspirated obstruents. We know that such forms existed in Ofo (for example, pop^{h}uti), so the constraint ranking established in (17)-(19), which was based upon the assumption that Ofo contained forms with identical aspirated obstruents, must be incorrect.

Given the constraints used here, one cannot construct a solution for the Ofo data set if it is understood to include identical aspirated obstruents. From this I conclude that forms with identical aspirated obstruents must have been prohibited in Ofo (recall that they are in fact absent in the Ofo lexical material available to us).

Let us assume, then, that the lack of Ofo forms with identical aspirated segments is not accidental, but rather that it represents a principled gap in the Ofo lexicon. In this case, the Ofo data differs from the Peruvian Aymara data in only two respects: first, identical aspirated obstruents do not cooccur; second, aspirated obstruents may cooccur with unaspirated but otherwise identical obstruents. The cooccurrence restrictions of Ofo can be generated by the constraint ranking shown in (21).

(21)

MAX-IO(seg) DEP-IO(seg) IDENTI→O(place) IDENTO→I(place) IDENTI→O(voice) IDENTO→I(voice) IDENTO→I(spread glottis) IDENTI→O(constricted glottis) IDENTO→I(constricted glottis)

↓

*SIM-4 *SIM-3 *SIM-2

↓

BEIDENTICAL IDENTI→O(spread glottis)

↓

*SIM-1

Notice that this ranking differs from that for Peruvian Aymara only in the position of BEIDENTICAL within the hierarchy (I am ignoring the positions of the IDENTI→O(constricted glottis) and IDENTO→I(constricted glottis) constraints, which are irrelevant, since there are no ejective segments in Ofo).

In tableaux (22)-(24), I show how the ranking in (21) will produce some crucial outputs: <u>haha</u> from /haha/, <u>t^hata</u> from /t^hat^ha/, and <u>t^hata</u> from /t^hata/.

(22)

	/haha/	IDI→O(place)	IDO→I(place)	IDO→I(sg)	*SIM-2	BEIDENTICAL	IDI→O(sg)
☞	haha						
	t^naha		*!				
	taha		*!				*

(23)

	/t^nat^na/	IDI→O(place)	IDO→I(place)	IDO→I(sg)	*SIM-2	BEIDENTICAL	IDI→O(sg)
	haha	**!			*	*	
	t^naha	*!				*	
☞	t^nata						*
	tata				*!		**

(24)

	/t^nata/	IDI→O(place)	IDO→I(place)	IDO→I(sg)	*SIM-2	BEIDENTICAL	IDI→O(sg)
	haha	**!					
☞	t^nata					*	
	tata				*!		*

5.2.3. Gojri

Gojri has voiced, voiceless unaspirated, and voiceless aspirated stops. The data for this language were discussed in 3.2.3. Gojri prohibits the cooccurrence of more than one aspirated stop per word, unless the aspirated stops are identical. Homorganic voiceless aspirated and voiceless unaspirated stops are also rare.

The Gojri cooccurrence restrictions can be generated by the constraint ranking shown in (25).

(25)

MAX-IO(seg)
DEP-IO(seg)
IDENTI→O(place)
IDENTO→I(place)
IDENTI→O(voice)
IDENTO→I(voice)
IDENTO→I(spread glottis)
IDENTI→O(constricted glottis)
IDENTO→I(constricted glottis)

↓

BEIDENTICAL

↓

*SIM-4
*SIM-3
*SIM-2

↓

IDENTI→O(spread glottis)

↓

*SIM-1

The tableaux in (26)-(28) illustrate the output candidate selection process for three input forms containing voiced stops (not found in Peruvian Aymara or Ofo).

(26)

	/dadi/	IDI→O(voice)	IDO→I(voice)	IDO→I(sg)	BEIDENTICAL	*SIM-2	IDI→O(sg)
☞	dadi						
	dati	*!					
	tati	**!					

(27)

	/dati/	IDI→O(voice)	IDO→I(voice)	IDO→I(sg)	BEIDENTICAL	*SIM-2	IDI→O(sg)
☞	dati						
	tati	*!					
	dadi		*!				

(28)

	/t^nadi/	IDI→O(voice)	IDO→I(voice)	IDO→I(sg)	BEIDENTICAL	*SIM-2	IDI→O(sg)
☞	t^nadi						
	t^nat^ni	*!					
	dadi		*!				*

5.3 Pattern Three

5.3.1 Bolivian Aymara

The cooccurrence restrictions of Bolivian Aymara are described in 3.3.1. Like Peruvian Aymara, this language has non-contrastive glottal stop, /h/, and voiceless unaspirated, voiceless aspirated, and ejective stops. The restrictions of this language differ from those of Peruvian Aymara in only two respects: first, heterorganic aspirated stops may cooccur; second, heterorganic aspirated and ejective stops may also cooccur. Only heterorganic ejectives, homorganic ejective and aspirated stops, glottal stop and ejectives, and homorganic aspirated or ejective and voiceless unaspirated segments are prohibited from cooccurring.

A ranking that will generate the restrictions of Bolivian Aymara is shown in (29).[3]

[3] This ranking abstracts away from ordering restrictions on the location of aspirated and ejective segments within the morpheme. See 6.2 for discussion.

(29)

MAX-IO(seg) DEP-IO(seg) IDENTI→O(place) IDENTO→I(place) IDENTI→O(voice) IDENTO→I(voice) IDENTO→I(spread glottis) IDENTO→I(constricted glottis)

↓

BEIDENTICAL

↓

*SIM-4 *SIM-3 IDENTI→O(spread glottis)

↓

IDENTI→O(constricted glottis)

↓

*SIM-2 *SIM-1

As noted above, Bolivian Aymara differs from Peruvian Aymara in just two respects: first, Bolivian Aymara allows aspirated and ejective stops to cooccur. The tableau in (30) shows that the input form /k'it^{h}a/ produces the output <u>k'it^{h}a</u>; this form involves a BEIDENTICAL violation, but nothing else. Identity (as in <u>k'ik'a</u> or <u>t^{h}it^{h}a</u>) cannot be achieved without violating IDENTI→O(place) and IDENTO→I(place) (not shown) – constraints that outrank BEIDENTICAL – so the BEIDENTICAL violation is tolerated.

(30)

	/k'it^{n}a/	BEIDENTICAL	*SIM-3	IDI→O(sg)	IDI→O(cg)	*SIM-2
☞	k'it^{n}a	*				*
	k'ita	*		*!		

Second, Bolivian Aymara allows heterorganic aspirated stops to cooccur, as shown in (31). Like the winning candidate in (31), p^{h}ut^{h}u involves only a BEIDENTICAL violation.

(31)

	/p^{n}ut^{n}u/	BEIDENTICAL	*SIM-3	IDI→O(sg)	IDI→O(cg)	*SIM-2
☞	p^{n}ut^{n}u	*				*
	p^{n}utu	*		*!		

The data of Bolivian Aymara is greatly complicated by ordering restrictions on the location of aspirated and ejective stops. These restrictions were described in 3.3.1.4; an analysis of those facts is offered in 6.2.

5.3.2 Hausa

Hausa has /h/, glottal stop, and three stop series: voiced, voiceless aspirated, and glottalic stop series. The glottalic series includes both implosive and ejective segments.

The restrictions of Hausa, described in 3.3.4, are repeated in (32).

(32)

a. Heterorganic glottalics (whether ejectives or implosives) cannot cooccur within roots.
b. Glottal stops do not cooccur with glottalics, whether ejectives or implosives.
c. Glottalic stops (whether ejective or implosive) tend not to cooccur with homorganic voiceless aspirated stops.

As noted in 3.3.4, restriction (32c) is subject to a directional asymmetry. Roots containing /tʰ/ followed by /ɗ/ seem rather common, while roots containing /ɗ/ followed by /tʰ/ are rare. I do not have an explanation for this asymmetry. The ranking presented in (33) generates the Hausa restrictions described above, except that /tʰ/⇨/ɗ/ is incorrectly ruled out.

(33)

MAX-IO(seg) DEP-IO(seg) IDENTI→O(place) IDENTO→I(place) IDENTI→O(voice) IDENTO→I(voice) IDENTO→I(constricted glottis) IDENTI→O(spread glottis)

↓

BEIDENTICAL

↓

*SIM-4 *SIM-3

↓

IDENTI→O(constricted glottis) IDENTO→I(spread glottis)

↓

*SIM-2 *SIM-1

Because the Hausa restrictions add nothing new to the data discussed so far in this chapter, I will not offer any tableaux for this language.

5.3.3. Old Georgian

Old Georgian had voiced, voiceless aspirated, and ejective stop series. The data from Old Georgian is discussed in 3.3.3; heterorganic ejective stops are prohibited from cooccurring within roots, as are homorganic ejective and aspirated stops.

The same constraint ranking given for Hausa will generate the cooccurrence restrictions of Old Georgian.

5.4 Pattern Four

5.4.1 Tzutujil

The Tzutujil data is described in 3.4.1. The language includes a glottalic series, composed of both ejective and implosive segments, a voiceless unaspirated plosive series, /h/, and glottal stop. Tzutujil has just two restrictions: one against heterorganic ejectives, and one against homorganic implosive or ejective and voiceless unaspirated stops. One ranking that will produce these restrictions is shown in (34).

(34)

MAX-IO(seg)
DEP-IO(seg)
IDENTI→O(place)
IDENTO→I(place)
IDENTO→I(spread glottis)
IDENTI→O(spread glottis)

↓

BEIDENTICAL

↓

*SIM-4
IDENTO→I(constricted glottis)

↓

IDENTI→O(constricted glottis)
IDENTI→O(voice)
IDENTO→I(voice)

↓

*SIM-3
*SIM-2
*SIM-1

As in earlier sections, not all constraints are shown in the following tableaux. Tzutujil differs from Pattern Three languages in that ejectives may cooccur with glottal stop, as shown in (35).

(35)

	/ʔak'/	MAX-IO(seg)	BEIDENTICAL	*SIM-4	IDO→I(cg)	IDI→O(cg)
☞	ʔak'		*			
	ʔa	*!				*
	ʔak		*			*!

The output candidates ʔaʔ and k'ak' (not listed) would involve violations of the undominated IDENTO→I(place) and IDENTI→O(place) constraints, so they could not be the winning output candidates.

Implosives can also cooccur with glottal stop; the relevant tableau is shown in (36).

(36)

	/ɓiiʔ/	MAX-IO(seg)	BE-IDENTICAL	*SIM-4	IDO→I(cg)	IDI→O(cg)	IDI→O (voice)	IDO→I (voice)
☞	ɓiiʔ		*					
	ɓii	*!						
	p'iiʔ		*				*!	

The analysis given here also predicts that implosives will fail to cooccur with voiceless homorganic stops. This appears to be true; there are no counterexamples in Dayley. However, note that /p/ and /t/ – the voiceless plosive counterparts of the two Tzutujil implosives, /ɓ/ and /ɗ/ – are rather uncommon in Tzutujil, so it is not clear just how robust this effect is. The tableau in (36) shows how such forms would be excluded.

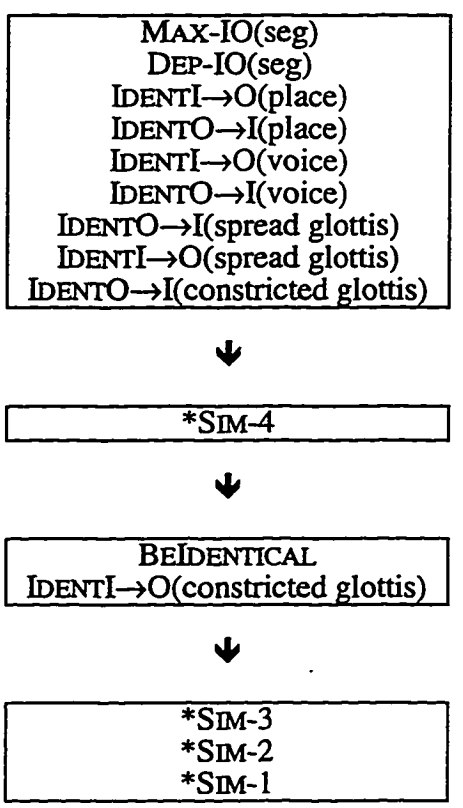
(37)

	/ɓap/	MAX-IO(seg)	BEIDEN-TICAL	*SIM-4	IDO→I(cg)	IDI→O(cg)	IDI→O (voice)	IDO→I (voice)
	ɓap		*!					
	p'ap		*!					
☞	pap							

5.4.1 Shuswap

The Shuswap data is described in 3.4.2. The language includes an ejective series and a voiceless unaspirated series, /h/, and glottal stop. The only restriction in Shuswap is that ejectives – whether identical or not – do not cooccur (recall that the language is characterized by the partial identity effect). Shuswap differs from Tzutujil in two respects: first, identical ejectives are not allowed to cooccur; second, ejectives and otherwise identical voiceless unaspirated stops *may* cooccur. One ranking that will produce these restrictions is shown in (38).

(38)



The tableau in (39) shows how identical ejectives are disallowed in Shuswap.

(39)

	/t'at'/	IDO→I(cg)	*SIM-4	BEIDENTICAL	IDI→O(cg)	*SIM-3
	t'at'		*!			
☞	t'at					
	tat		*!			

The tableau in (40) shows how input forms with ejectives and otherwise identical voiceless unaspirated stops can surface as such.

(40)

	/k'ik/	IDO→I(cg)	*SIM-4	BEIDENTICAL	IDI→O(cg)	*SIM-3
☞	k'ik					
	kik		*!			
	k'ik'	*!				

5.5 Conclusion

The laryngeal cooccurrence restrictions of the languages discussed here prevent the cooccurrence of similar elements. I have analyzed these cooccurrence restrictions with reference to a family of *SIMILARITY constraints that proscribe a certain degree of similarity between pairs of elements occurring within morphemes.

There is also an identity effect found in some languages (for example, Peruvian Aymara, Gojri, Bolivian Aymara, Hausa, Old Georgian, and Tzutujil). Two /h/s, identical aspirated stops, identical glottal stops, and identical ejectives may cooccur (subject to inventory considerations and restrictions on the location of glottal stop) in these languages; they do not cooccur in Cuzco Quechua, Souletin Basque, or Sanskrit.

Identical elements may cooccur when BEIDENTICAL outranks all *SIMILARITY constraints. The formulation of BEIDENTICAL as a single constraint predicts that identical elements will often be allowed to cooccur as a block: if a language allows t^{h}at^{h}a, for example, then (inventory considerations aside) it will also allow t'at'a, haha, and ʔaʔa.

Shuswap and, I have argued, Ofo, illustrate the partial identity effect. In this case, two /h/s or two glottal stops may cooccur within morphemes, but identical aspirated or ejective elements with supralaryngeal place features cannot cooccur. These restrictions result when the active *SIMILARITY constraint outranks BEIDENTICAL and some IDENTI→O(laryngeal feature) constraint(s) but no IDENTO→I(laryngeal feature) constraints. No constraint ranking would generate the opposite case, in which identical ejectives could cooccur but two glottal stops, for example, could not. I know of no examples of languages with such phenomena.

Finally, when BEIDENTICAL outranks all *SIMILARITY constraints, and these in turn outrank some subset of the laryngeal Correspondence constraints, the resulting grammar will generate a language that will sacrifice (or insert) a laryngeal feature in order to ensure identical consonants. In other words, there will be no morphemes with homorganic aspirated and plain stops, or with homorganic ejective and plain stops (*t^{h}ata, *t'ata). This accounts for the striking fact that the lack of these forms is the sole respect in which the cooccurrence restrictions of languages with the complete identity effect – most of the languages in Patterns Two, Three, and Four – are stricter than those of the languages in Pattern One.

CHAPTER 6: ADDITIONAL ANALYSIS

Data from eleven languages were presented in chapter 3. Four of these languages[1] involve complexities not accounted for by the analyses presented in the preceding two chapters. Cuzco Quechua, Souletin Basque, Peruvian Aymara, and Bolivian Aymara all involve a leftness effect ([constricted glottis] and [spread glottis] features occur as close to the beginnings of morphemes as possible) that has not yet been addressed. This effect is discussed in the following section.

The presence in Bolivian Aymara of forms involving both aspirated and ejective stops gives rise to an interesting set of ordering facts not found in other languages with the leftness effect. An analysis of this data is offered in 6.2.

6.1 The Leftness Effect in Cuzco Quechua, Souletin Basque, and Peruvian Aymara

Recall from 3.1.1 that Cuzco Quechua, a Pattern One language, has three stop series: voiceless unaspirated, voiceless aspirated, and ejective. This language allows only one aspirated or ejective stop per morpheme. If an aspirated or ejective stop is present, it will be the first stop in the morpheme: t'anta 'bread', sut'a 'volcanic rock or lava', *tant'a. In other words, no voiceless unaspirated stop can precede an ejective or an aspirated stop.

[1] Actually five, but I do not treat the Sanskrit alternations in this study.

Souletin Basque, likewise a Pattern One language, is discussed in 3.1.2. This language has three stop series: voiceless unaspirated, voiceless aspirated, and voiced. Again, only one aspirated stop can be present per morpheme, and such a stop will be the first voiceless stop in the morpheme: k^{h}ako 'to hook', gark^{h}otʃ 'nape', *kak^{h}o. No voiceless unaspirated stop can precede an aspirated stop.

Peruvian Aymara (see 3.2.1) has a segment inventory that is virtually identical to that of Cuzco Quechua. In this language, as in Cuzco Quechua, if an ejective or aspirated stop is present, then it will be the first stop in the morpheme: k^{h}antʃa 'tuber damaged during the harvest', wak^{h}a 'untamable', *kantʃ^{h}a. (In Peruvian Aymara, however, identical aspirated stops, and identical ejectives, are allowed to cooccur.) No voiceless unaspirated stop can precede an ejective or aspirated stop.

The simple leftness effects of Cuzco Quechua, Souletin Basque, and Peruvian Aymara can be handled with the constraints given in chapter 4 plus two new ones. These constraints, LEFTMOST(spread glottis) and LEFTMOST(constricted glottis), are given in (1).

(1)

1. **LEFTMOST(spread glottis) – Aspiration features should occur early in the morpheme.** One violation is assessed for every available host consonant intervening between the beginning of the morpheme and the location of aspiration.

2. **LEFTMOST(constricted glottis) – Ejective features should occur early in the morpheme.** One violation is assessed for every available host consonant intervening between the beginning of the morpheme and the location of the ejective feature.

The LEFTMOST constraints require laryngeal features to be near the leading edges of morphemes. One violation is counted for every available host segment intervening between the [spread glottis] or [constricted glottis] feature and the beginning of the morpheme. The forms tat^{h}a, tat'a, and taha would each incur two LEFTMOST violations; sap'a, t'ant'a, and t'anta would earn none.[2]

There are at least two possible bases for the LEFTMOST constraints. First, restricting any phenomena to the beginnings or ends of words (or, incidentally, to roots as opposed to affixes) gives hearers additional cues to word delimitation (Trubetzkoy 1969). It is in this spirit that Carenko (1975:14) reports on ". . . the delimitative function of laryngealization in Quechua: segmental aspiration and glottal occlusion serve as signals for the beginning of a word."

Second, restricting contrasts to a particular position in a word reduces ambiguity and lessens the hearer's burden of recognition. If the position of expanded contrasts comes early in the word, then the hearer has the added advantage of getting more valuable information – information that distinguishes among more lexical items – earlier.

The point of uniqueness of a word is the moment at which it can be distinguished from all other forms in the hearer's lexicon. If one accepts the forms in Hornberger and

[2] In MacEachern 1996, an analysis of Bolivian Aymara, LEFTMOST violations were counted according to the number of segments – of any type – intervening between a [spread glottis] or [constricted glottis] feature and the beginning of the morpheme. According to that formulation of the constraint, t'ata earned no LEFTMOST violations, while sap'a and t'ant'a earned two and three, respectively.

Hornberger 1983 as a model of the Cuzco Quechua lexicon, for example, then the point of uniqueness of p'isaqa ('partridge') occurs when the /a/ is heard. No other forms in the dictionary begin with p'isa-, so it is only at this point that p'isaqa becomes distinct from, for example, p'isqaka 'golden-billed saltator' and p'istukuj 'cover up, wrap oneself up'. Extensive psycholinguistic evidence exists showing that lexical decisions are made faster when the point of uniqueness lies earlier in a word (Marslen-Wilson and Welsh 1978, Marslen-Wilson and Tyler 1980, Marslen-Wilson 1984, 1987, Emmorey 1987).

Notice that the possible constraints RIGHTMOST(spread glottis) and RIGHTMOST(constricted glottis) would serve equally well as boundary signals. However, if ejective and aspiration features were subject to RIGHTMOST rather than LEFTMOST, then the position of greatest consonantal contrasts would come at the trailing edge of the morpheme. The point of uniqueness of any given morpheme would tend to lie farther to the right, and lexical decisions would be slower.

A constraint ranking that will generate the leftness effect of Cuzco Quechua is shown in (2). Note that apart from the presence of LEFTMOST(spread glottis) and LEFTMOST(constricted glottis), the ranking is the same as that provided in 5.1.1.

(2)

MAX-IO(seg)
DEP-IO(seg)
IDENTI→O(place)
IDENTI→O(voice)
IDENTO→I(voice)
IDENTO→I(spread glottis)
IDENTO→I(constricted glottis)
LEFTMOST(constricted glottis)

↓

LEFTMOST(spread glottis)

↓

*SIM-4
*SIM-3
*SIM-2
*SIM-1

↓

BEIDENTICAL
IDENTI→O(spread glottis)
IDENTO→I(place)
IDENTI→O(constricted glottis)

Sample tableaux are given in (3)-(5). In (3) below, an input form with a voiceless unaspirated stop followed by a heterorganic ejective maps onto an output composed of two voiceless unaspirated stops.

(3)

	/kat'a/	IDO→I(cg)	LEFT(cg)	LEFT(sg)	*SIM-1	IDI→O(cg)	BEIDENTICAL
	kat'a		**!				
☞	kata						
	k'at'a	*!					

In tableau (4), we find a similar result for homorganic segments.

(4)

	/tat'a/	IDO→I(cg)	LEFT(cg)	LEFT(sg)	*SIM-1	IDI→O(cg)	BEIDENTICAL
	tat'a		**!				*
☞	tata					*	
	t'at'a	*!					

Finally, (5) illustrates why LEFTMOST(spread glottis) cannot be undominated in Cuzco Quechua.

(5)

	/taha/	IDO→I(sg)	IDO→I(cg)	LEFT(cg)	LEFT(sg)	*SIM-1	BEIDENTICAL
☞	taha						*
	t'aha		*!				*
	t^{h}aha	*!					*

If LEFTMOST(spread glottis) were undominated, then t'aha would emerge the winner, and we would be unable to account for the presence of forms such as taha in Cuzco Quechua.

Regarding Souletin Basque, the leftness effect in that language can be generated from the constraint ranking shown in (6). Notice that this ranking is the same as that given in 5.1.2, except for the addition of the two LEFTMOST constraints (LEFTMOST(constricted glottis) is actually irrelevant to Souletin Basque, since the language does not contain ejectives or glottal stop).

(6)

MAX-IO(seg)
DEP-IO(seg)
IDENTI→O(place)
IDENTI→O(voice)
IDENTO→I(voice)
IDENTO→I(spread glottis)
IDENTI→O(constricted glottis)
IDENTO→I(constricted glottis)
LEFTMOST(spread glottis)
LEFTMOST(constricted glottis)

↓

*SIM-4
*SIM-3
*SIM-2
*SIM-1

↓

BEIDENTICAL
IDENTI→O(spread glottis)
IDENTO→I(place)

Two tableaux illustrating the leftness effect of Souletin Basque are shown below.

(7)

	/kat^{n}a/	IDI→O(pl)	IDO→I(sg)	LEFT(sg)	*SIM-1	BEIDENTICAL	IDI→O(sg)
	k^{n}ata		*!				
☞	kata						
	kat^{n}a			**!			
	kaha	*!					

(8)

	/tat^{n}a/	IDI→O(pl)	IDO→I(sg)	LEFT(sg)	*SIM-1	BEIDENTICAL	IDI→O(sg)
	t^{n}ata		*!				
☞	tata						
	tat^{n}a			**!			
	taha	*!					

A constraint that will generate the leftness effect of Peruvian Aymara is shown in (9). Again, note that apart from the insertion of the two LEFTMOST constraints, the ranking is the same as that provided in 5.2.1.

(9)

MAX-IO(seg) DEP-IO(seg) IDENTI→O(place) IDENTO→I(place) IDENTI→O(voice) IDENTO→I(voice) IDENTO→I(spread glottis) IDENTO→I(constricted glottis) LEFTMOST(spread glottis) LEFTMOST(constricted glottis)

↓

BEIDENTICAL

↓

*SIM-4 *SIM-3 *SIM-2

↓

IDENTI→O(constricted glottis) IDENTI→O(spread glottis)

↓

*SIM-1

I do not present any tableaux from Peruvian Aymara here, since the language adds nothing of interest to what has already been discussed.

6.2 The Ordering Restrictions of Bolivian Aymara

As mentioned above, Bolivian Aymara has a more complicated set of ordering restrictions than those found in the languages discussed in the preceding section; this data was covered in 3.3.1.4. If there is a single ejective or aspirated stop in a morpheme, then that stop will be the first stop in the morpheme. In this respect, the language is the same as Cuzco Quechua and Peruvian Aymara. No voiceless unaspirated stop can precede an ejective or aspirated stop.

However, there are some Bolivian Aymara forms that contain both an ejective and an aspirated stop. In these forms, the ejective must precede the aspirated stop (t'alp^{h}a 'wide', k'ip^{h}a 'said of late potatoes'), unless the first stop in the morpheme is a bilabial or uvular. In that case, the aspirated stop will precede the ejective stop (p^{h}ant'a 'black coat', q^{h}ot'a 'resin of some small plants').

The ordering restrictions of Bolivian Aymara require reference to five new constraints, in addition to the constraints from chapter 4, and the two LEFTMOST constraints introduced in the preceding section. The five new constraints, PRESERVE(laryngeal feature), *h⇔ʔ, AVOID p', and EJECTIVES PRECEDE ASPIRATES, are set out in (10) below.

(10)

1. **PRESERVE(laryngeal feature) – Laryngeal features present in the input must also be present in the output.** One violation is assessed for every laryngeal feature present in the input that is not present in the output.

2. ***h⇔ʔ – /h/ and /ʔ/ cannot be correspondents.** One violation is assessed for every instance in which /h/ corresponds to /ʔ/, or vice-versa.

3. **AVOID p' – Avoid bilabial ejectives.** One violation is assessed for every bilabial ejective present in the output.

4. **AVOID q' – Avoid uvular ejectives.** One violation is assessed for every uvular ejective present in the output.

5. **EJECTIVES PRECEDE ASPIRATES (E/A)– Ejective stops must precede aspirated stops.** One violation is assessed for every instance in which an ejective stop follows an aspirated stop.

Regarding the PRESERVE(laryngeal feature) constraint, notice that the location of ejective and aspiration features in Bolivian Aymara is predictable. This suggests that there is a value attached to having features present in the input also be present in the output, even if those features reside on different segments in the output – in other words, even when violations of the IDENTI→O(laryngeal feature) and IDENTO→I(laryngeal feature) constraints are involved. No provisions were made for this effect in McCarthy and Prince 1995, but the authors (p. 265) recognize that such an extension is necessary (for example, in order to handle floating features). The PRESERVE(laryngeal feature) constraint achieves this end. The input-output pair /tat'a/-tata would involve one PRESERVE(laryngeal feature) violation; the input-output pair /tat'a/-t'ata would involve no PRESERVE(laryngeal feature) violations, although the form would of course violate IDENTI→O(constricted glottis) and IDENTO→I(constricted glottis).

The constraint *h⇔ʔ penalizes input-output correspondence between /h/ and /ʔ/.

AVOID p' prohibits bilabial stops from being ejective. The output form p'ata would involve one violation of this constraint, while the output form p'ap'a would involve two violations.

The AVOID q' constraint penalizes uvular ejectives; violations are counted as for AVOID p'. Possible physical bases for the existence of these constraints are suggested in 3.3.1.4.

Finally, EJECTIVES PRECEDE ASPIRATES penalizes any form in which an aspirated stop precedes an ejective stop. For example, t^hak'i would garner one E/A violation; t'ak^{h}i would incur no E/A violations. Some constraint of this type is necessary to account for the difference in wellformedness between Bolivian Aymara forms such as t'ink^{h}a 'tip' and *t^hink'a (unattested).[3,4]

One constraint ranking that will generate the ordering restrictions of Bolivian Aymara is shown in (11). Note that the constraint ranking is quite different from that provided in 5.3.1, which produced the cooccurrence restrictions of Bolivian Aymara but abstracted away from the ordering restrictions of this language.

---

[3] In earlier work on Bolivian Aymara (MacEachern 1996), I analyzed this effect by appealing to two LEFTMOST constraints. One constraint penalized ejective features that were not at the leading edge of the morpheme, and a second constraint penalized *all* laryngeal features that were not at the leading edge of the morpheme. As noted in footnote 2 of chapter 6, violations of those LEFTMOST constraints were not calculated in the same way as violations of the LEFTMOST constraints discussed here.

[4] Whether or not EJECTIVES PRECEDE ASPIRATES and *h⇔ʔ will be useful in analyses of other languages remains to be seen.

(11)

MAX-IO(seg)
DEP-IO(seg)
IDENTI→O(place)
IDENTO→I(place)
IDENTI→O(voice)
IDENTO→I(voice)
*h⇔ʔ
LEFTMOST(constricted glottis)

↓

BEIDENTICAL

↓

*SIM-4
*SIM-3

↓

PRESERVE(lar)

↓

AVOID p'

↓

AVOID q'

↓

E/A

↓

IDENTI→O(spread glottis)

↓

IDENTO→I(spread glottis)

↓

IDENTI→O(constricted glottis)

↓

IDENTO→I(constricted glottis)

↓

LEFTMOST(spread glottis)

↓

*SIM-2
*SIM-1

Sample tableaux for several possible inputs are given in (12)-(20). The need for *h⇔ʔ can be seen by inspecting the tableau in (12). If not for this constraint, input forms containing glottal stop and /h/ would surface with two /h/s (recall that /ʔ/ does not occur in medial position in Bolivian Aymara; in any case, ʔaʔa would be less optimal than haha due to an IDENTI→O(spread glottis) violation).

(12)

	/ʔaha/	*h⇔ʔ	BEIDENTICAL	*SIM-3	PRES(lar)	IDO→I(sg)	IDI→O(cg)
☞	ʔaha						
	haha	*!		*	*		*

Tableau (13) illustrates the cooccurrence of an ejective stop with an aspirated stop. The faithful form emerges the winner.

(13)

	/t'ak^{n}a/	BEIDENTICAL	PRES(lar)	E/A	IDI→O(sg)	IDI→O(cg)	IDO→I(cg)
☞	t'ak^{n}a	*					
	t^{n}ak'a	*		*!			
	t'aka	*	*!		*	*	

Tableau (14) shows how a hypothetical input, containing an aspirated and an ejective stop in the "wrong" order, might be fixed by the grammar of the language. This tableau highlights the importance of PRESERVE(laryngeal feature) and EJECTIVES PRECEDE ASPIRATES.

(14)

	/t^nak'a/	BEIDENTICAL	PRES(lar)	E/A	IDI→O(sg)	IDI→O(cg)	IDO→I(cg)
	t^nak'a	*		*!			
☞	t'ak^{n}a	*					*
	t^naka	*	*!				

Tableau (15) establishes that a bilabial aspirated stop can surface with a following ejective.

(15)

	/p^nat'a/	BEIDENTICAL	PRES(lar)	AVOID p'	E/A	IDI→O(sg)	IDO→I(sg)
☞	p^nat'a	*					
	p'at^{n}a	*		*!			
	p^nata	*	*!				

Tableau (16) assumes a hypothetical input form in order to establish that when the first stop in the word is bilabial, the aspirated stop must precede the ejective stop. (Outcomes analogous to those shown in (15) and (16) will result when an initial uvular stop is paired with a stop at an intermediate place of articulation, and the morpheme contains one aspiration feature and one ejective feature).

(16)

	/p^{n}at'a/	BEIDENTICAL	PRES(lar)	AVOID p'	E/A	IDI→O(sg)	IDO→I(sg)
	p'at^{n}a	*		*!			*
☞	p^{n}at'a	*					
	p^{n}ata	*	*!				

The tableau in (17) shows that ejective bilabials *can* occur in Bolivian Aymara, despite the presence of the constraint AVOID p'. A similar result is produced for uvular ejectives.

(17)

	/p'aka/	LEFTMOST(cg)	BEIDENTICAL	PRES(lar)	AVOID p'	IDI→O(cg)	IDO→I(cg)
☞	p'aka		*				
	paka		*	*!			
	pak'a	**!					*

My dictionary source for Bolivian Aymara (De Lucca 1987) lists only two forms from the standard dialect that contain a uvular stop, a bilabial stop, and an aspiration feature and an ejective feature. These forms are q'ap^{h}i 'fragrance' and q'ap^{h}a 'active, diligent, agile'. The ranking given in (11) will produce these forms from the inputs /q'ap^{h}i/ and /q'ap^{h}a/. I noted in 3.3.1.4 that forms containing initial /q^{h}/ and medial /p'/ are present in other dialects (i.e., q^{h}op'aki 'meal with meat or fat' from the Altiplano Central dialect, and q^{h}op'i 'potter' from the Altiplano Norte dialect). As (18) shows, the ranking given in (11) makes the prediction that such forms would emerge with initial uvular ejectives and medial aspirates in the standard dialect.

(18)

	/q^{n}op'i/	BEIDENTICAL	*SIM-3	PRES(lar)	AVOID p'
	q^{n}op'i	*			*!
☞	q'op^{n}i	*			
	q^{n}opi	*		*!	

Finally, forms containing initial bilabial stops, medial uvular stops, and one aspiration feature and one ejective feature (not found in De Lucca) are all predicted to emerge as *p^h...q'* sequences. This can be seen from tableaux (19) and (20).

(19)

	/p^{n}oq'i/	BEIDENTICAL	*SIM -3	PRES(lar)	AVOID p'
☞	p^{n}oq'i	*			
	p'oq^{n}i	*			*!
	p^{n}oqi	*		*!	

(20)

	/p'oq^{n}i/	BEIDENTICAL	*SIM-3	PRES(lar)	AVOID p'
	p'oq^{n}i	*			*!
☞	p^{n}oq'i	*			
	p'oqi	*		*!	

This concludes my account of the ordering restrictions of Bolivian Aymara.

CHAPTER 7: OTHER APPROACHES

7.1 Traditional OCP-based approaches

The Obligatory Contour Principle (OCP) was proposed by Leben in 1973. As originally conceived, it was formulated as a universal principle that prohibited adjacent, identical tonemes. The principle was named and elaborated upon by Goldsmith 1976.

McCarthy 1979 and McCarthy 1981 were major early works that extended the OCP beyond tonal phenomena. McCarthy treated the OCP as a universal condition that applied to all autosegmental features. In his research, the OCP assumed the following shape: "Adjacent identical elements are prohibited."

The OCP has served as the primary ingredient in Feature Geometric analyses of cooccurrence restrictions in many studies, including Hualde 1993, Ito and Mester 1986, Lombardi 1991, McCarthy 1986, 1988, 1989, Mester 1986, Padgett 1992, 1995, Parker and Weber 1996, Shaw 1991, and Yip 1988, 1989. OCP-based accounts of laryngeal cooccurrence restrictions at first appear promising. The existence of an independent Laryngeal node combines with the theory of underspecification – only some sounds are distinctively specified for laryngeal features – to predict the existence of cooccurrence restrictions operating on [voice], [spread glottis], [constricted glottis], or the Laryngeal node itself (see Shaw 1991 for more extensive discussion of this point). In this section, however, I suggest that OCP-based analyses fail to capture the data presented in chapter 3. Arguments that traditional OCP-based analyses cannot account for other types of

cooccurrence restrictions have been made by other researchers: Pierrehumbert 1993, Frisch, Broe and Pierrehumbert 1995, and Frisch 1996 offer details on how such analyses fail to account for data on the cooccurrence (and non-cooccurrence) of homorganic and identical consonants in Arabic verbal roots.

Two particular shortcomings of the OCP are itemized below.

1. Featural identity is a poor way of describing different degrees of similarity, whether the gradiency at issue is the strength of the restriction (as in the Arabic place restrictions, described in various accounts; see Frisch 1996 for analysis and references), or the graded nature of similar categorical restrictions in different languages (as in Cuzco Quechua and the Aymara dialects). OCP-based analyses also do not capture cross-linguistic regularities in the patterning of cooccurrence restrictions; these regularities appear to be based on relative degrees of similarity.

2. There is difficulty in accounting for the identity effect without violating theory-internal principles. OCP-based analyses allow the identity effect through a branching Root node device; this device does not extend to cover the Aymara data, where consonants can intervene between identical elements.

In 7.1.1, I present a summary of McCarthy 1989, an OCP-based analysis of cooccurrence restrictions in the Mayan languages. This account is representative of many

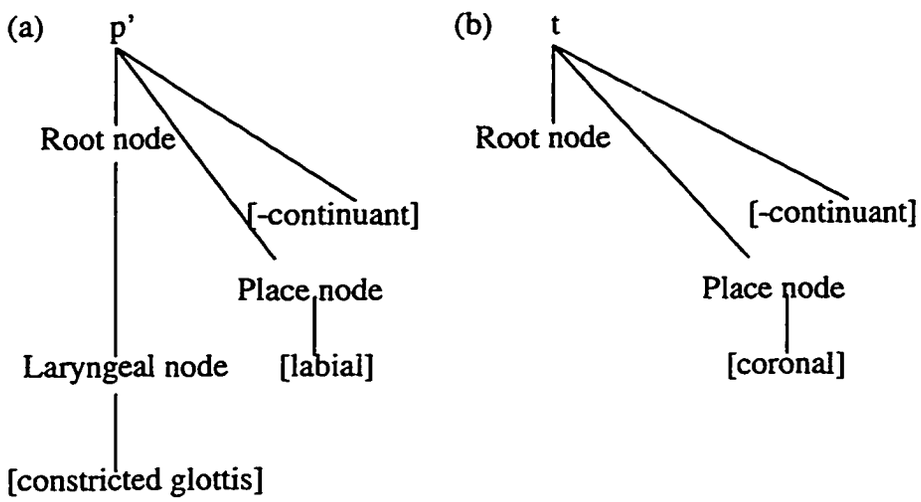
OCP-based Feature Geometric accounts of cooccurrence restrictions published in the past decade. In 7.1.2 and 7.1.3, I address the points mentioned above, and seek to establish that the OCP is fundamentally flawed in these respects. In 7.1.4, I present a brief summary of the conclusions reached here.

7.1.1 McCarthy 1989

McCarthy 1989 offers an analysis of the Mayan prohibition on heterorganic ejectives (specifically, restrictions in the Tsotsil, Chontal, Yucatec, and Tzutujil languages). The author attempts to account for only one of the restrictions discussed in 3.4.1, namely that which prohibits heterorganic ejectives from cooccurring in Mayan roots (*t'ak'). Recall that identical ejectives *are* allowed to cooccur (for example, tʃ'ihtʃ' 'metal, car' and q'iiq' 'north wind' are found in Tzutujil).

In McCarthy's analysis, voiceless unaspirated consonants lack a Laryngeal node, while ejective consonants are characterized by a [constricted glottis] specification on the Laryngeal node. This produces structures like those shown in (1). (Throughout this and following sections, irrelevant details of the representations are omitted; the particular version of Feature Geometry used here is not crucial to the argument.)

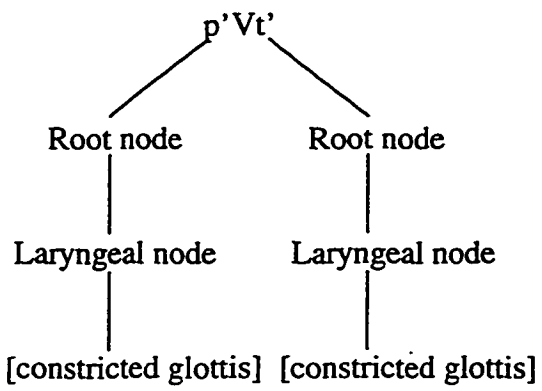
(1)



McCarthy's analysis requires that vowels, which may intervene between identical ejectives, be transparent to consonant association lines. This state of affairs is brought about through C-V planar segregation, first argued for in McCarthy 1979. Given the typically CVC structure of Mayan roots, McCarthy suggested that the order of consonants and vowels in these roots is redundant. Consonants and vowels reside on separate tiers, and therefore vowel features do not intervene between consonant features.

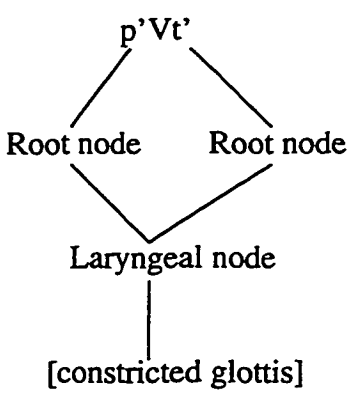
Adjacent, identical Laryngeal nodes are prohibited, as dictated by the OCP. When combined with C-V planar segregation, this constraint will rule out structures like the one shown in (2).

(2)



Heterorganic ejectives can still occur when one Laryngeal node links to two Root nodes, as shown in (3).

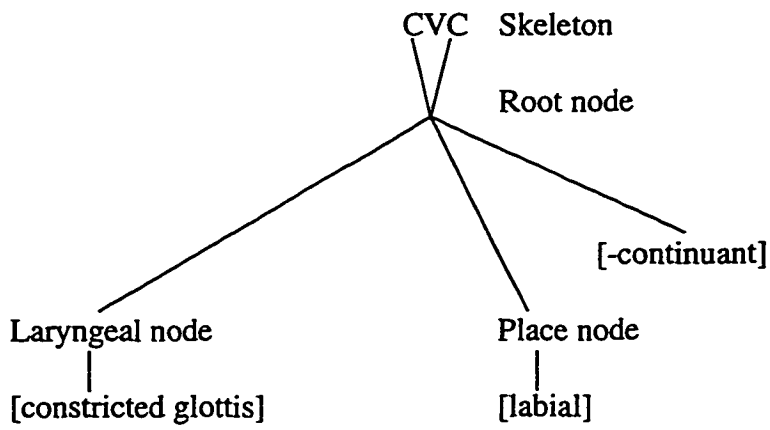
(3)



The structure in (3) can be disallowed through a language-particular constraint prohibiting Laryngeal nodes from linking to multiple Root nodes. The Mayan languages, McCarthy suggested, have this constraint. Therefore, heterorganic ejectives do not occur in these languages. Identical ejectives, however, can occur when the Root node branches

(see Mester 1986 for an earlier analysis making use of branching Root nodes). The root p'Vp' would be diagrammed as shown in (4).

(4)



Language-specific constraints prohibiting branching Root nodes can exist. Under this view, it is the existence of such constraints that accounts for languages (like Cuzco Quechua) that prohibit ejectives, whether identical or not, from cooccurring within morphemes.

To sum up, McCarthy 1989 accounts for the Mayan prohibition against heterorganic ejectives through reference to an OCP ban on adjacent identical Laryngeal nodes (specifically, adjacent [constricted glottis] specifications). Most languages allow such configurations by linking a single Laryngeal node to multiple Root nodes. In the Mayan languages, however, a language-particular constraint prohibits the double-linking of Laryngeal nodes. Identical ejectives *are* found, however; they occur when Root nodes branch. In languages with morpheme structure constraints that disallow the cooccurrence

of two ejectives, regardless of whether or not they are identical, branching Root nodes are disallowed.

Vowels may intervene between consonants instantiating branching nodes because, in the Mayan languages, they exist on a separate plane from consonants.

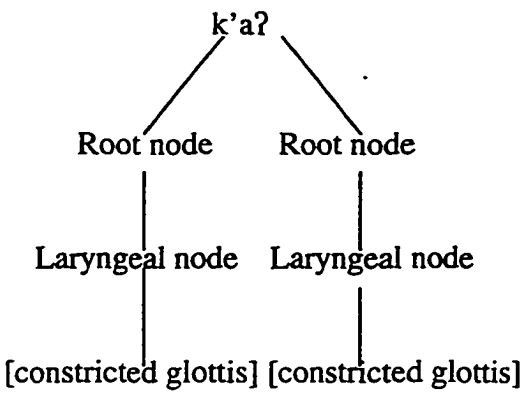
7.1.2 The OCP and similarity

The analysis summarized in 7.1.1 does not account for the nonparticipation of glottal stop and implosives in the Mayan restrictions against heterorganic ejectives, which are characterized as affecting [constricted glottis] specifications. Both glottal stop and implosives are commonly understood to involve [constricted glottis], so we would expect them to be prohibited from cooccurring with ejectives and with each other. They are not.

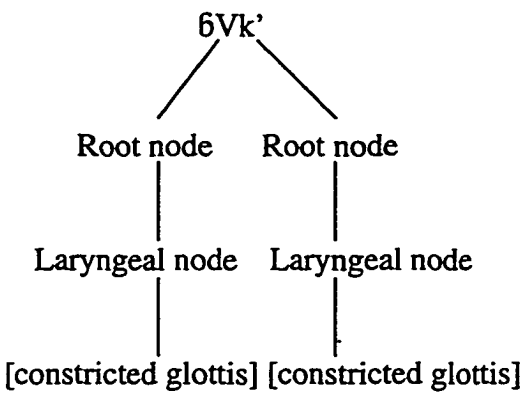
McCarthy 1989 considers only the voiceless pulmonic series and the ejective series. That article suggests that implosives may be derived, and that this is the reason they do not participate in the cooccurrence restrictions. The article does not discuss the absence of cooccurrence restrictions involving glottal stop. Consider (5a) and (5b): the segments involved are not identical, so they cannot be connected by branching Root nodes. The analysis recounted in the previous section argues for language-specific constraints prohibiting a single [constricted glottis] specification from branching to non-identical segments. Each root, then, must contain two [constricted glottis] specifications.

(5)

(a)



(b)



The existence of identical, adjacent [constricted glottis] specifications wrongly suggests that forms containing an implosive and a glottal stop, or an ejective and a glottal stop, or an implosive and an ejective, violate the OCP and should therefore be ruled out.

Notice, with respect to the material presented in 3.4.1, that it is the cooccurrence of two elements with [constricted glottis] specifications that are holistically more similar – i.e. two ejectives – that is prohibited. The cooccurrence of two elements with [constricted glottis] specifications that are less similar – for example, glottal stop and an

implosive, or glottal stop and an ejective, or an implosive and an ejective – is not prohibited.

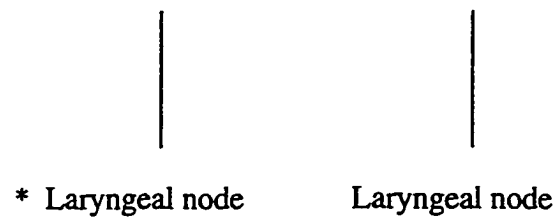
In more general terms, the analysis presented in 7.1.1 does not capture the gradient similarity effects seen in action throughout chapter 3.[1] Consider Cuzco Quechua. We might suggest that the language disallows two [constricted glottis] specifications within a morpheme. This would account for the prohibition against having more than one ejective per morpheme. We might also suggest that the language disallows two [spread glottis] specifications (indicating aspiration) within a morpheme, in order to account for the lack of morphemes with two /h/s or two aspirated stops, or an /h/ and an aspirated stop.

However, in order to account for the ban on the cooccurrence of ejective and aspirated consonants, we would need to prohibit [constricted glottis] from cooccurring with [spread glottis]. We might do this by proposing that roots cannot contain two distinct Laryngeal node specifications, as shown in (6). Such a restriction would prohibit

[1] Yip 1989 follows McCarthy in his analysis of the cooccurrence restrictions operating in the Mayan languages. Yip's article, however, expresses ideas similar to mine regarding the relevance of gradient similarity. One statement in her article prefigures the account of cooccurrence restrictions presented in this dissertation: "Some notion of relative similarity or partial identity must be used to capture the intuition that dentals are the quintessential coronals of Javanese and thus cannot occur with any others, whereas the retroflexes and palatals are more different from each other than each is from the dentals, and thus can cooccur." This statement broaches the shortcomings of the traditional OCP with respect to restrictions on the occurrence of identical feature specifications, which do not lend themselves to capturing the relevant similarity continuums.

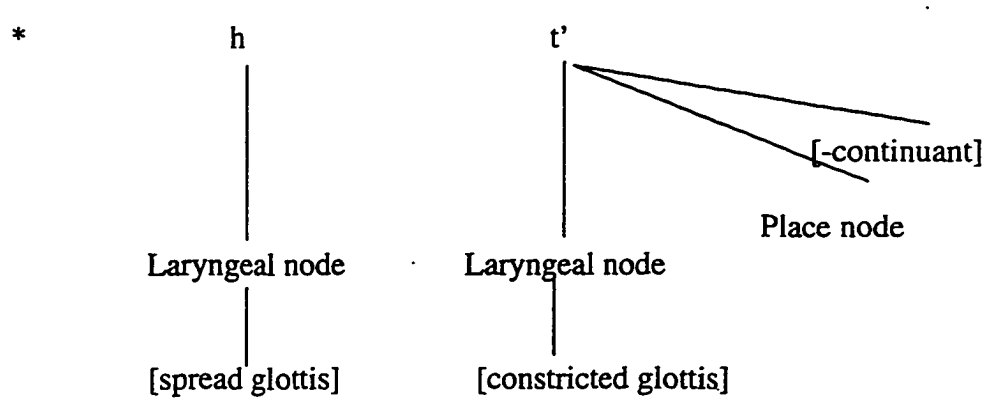
neighboring specifications for [spread glottis] and [spread glottis], or [constricted glottis] and [constricted glottis], or [spread glottis] and [constricted glottis].

(6)



While this analysis would correctly prohibit forms containing two aspirated stops, two ejectives, two /h/s, an ejective and an aspirated stop, and an /h/ and an aspirated stop, it would incorrectly rule out the occurrence of /h/ with ejective consonants, as shown in (7).

(7)



What is basic to the failure of this analysis is that it does not capture the notion of a similarity continuum. The similarity continuum suggests that /h/ and aspirated stops are

too similar to cooccur, and aspirated stops and ejectives are too similar to cooccur, but /h/ and ejectives are *not* too similar to cooccur.

Of course, an analysis of the type described in 7.1.1 could be made to work. In order to do so, however, we would need to complicate the account by incorporating three separate prohibitions, one of which would have to make reference to whether or not a Laryngeal node specification is linked to a [-continuant] feature. A sample set of conditions that would accomplish the feat is shown in (8).

(8)

(a) A morpheme may not contain two [spread glottis] specifications.
(b) A morpheme may not contain two [constricted glottis] specifications.
(c) A morpheme may not contain two segments with Laryngeal node specifications that are also [-continuant] (this statement assumes that glottal stop is not [-continuant]).

The condition listed in (8a) would rule out combinations of two aspirated stops, two /h/s, or an /h/ and an aspirated stop. The statement in (8b) would rule out combinations of two ejectives, and combinations of ejectives and glottal stop. (Recall that glottal stop is only found in initial position in Cuzco Quechua roots, so no prohibition is needed on the occurrence of two glottal stops.) Finally, (8c) is necessary to prevent the cooccurrence of aspirated stops with ejectives; it will also, of course, overlap the domains of (8a) and (8b) by prohibiting two aspirated stops and two ejectives.

This type of analysis of the Cuzco Quechua data, although workable, does not give an insightful picture of the cooccurrence restrictions present in that language.

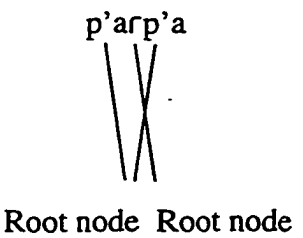
It is also the case that OCP-based analyses do not clarify the typology of laryngeal cooccurrence restrictions. Consider Table 3.5. There are no cases of languages that prohibit two elements characterized by [constricted glottis], and do not also prohibit two elements characterized by [spread glottis]. OCP-based accounts would lead one (incorrectly, it seems) to expect that such restrictions will operate independently of one another.

7.1.3 The OCP and the identity effect[2]

Although the analysis presented in 7.1.1 did account for the identity effect in the Mayan languages, the existence of the identity effect in Aymara proves problematic for this account. In 7.1.1, the presence of identical ejectives within roots was attributed to C-V planar segregation and branching Root nodes. First, it is not clear that C-V planar segregation could be argued for in Aymara. Setting aside the question of the relationship between vowels and consonants, however, it is still clear that the branching Root node device cannot serve for Aymara, where consonants may intervene between identical ejectives. The presence of such consonants would entail violations of the prohibition against crossed association lines (Goldsmith 1976). For example, p'arp'a ('potter's clay' in Bolivian Aymara) would be diagrammed as shown in (9).

[2] See also Pierrehumbert 1993, which details the failure of the OCP to properly account for the lack of Arabic verbal roots with non-adjacent identical segments (i.e. mdm).

(9)



This geometry is illegal, due to the crossed association lines. It is not clear how OCP-based Feature Geometric analyses could be extended to account for the Aymara data.

7.1.4 Section summary

I conclude that traditional OCP-based analyses do not extend to cover the data discussed in this dissertation.

Recently, however, implementations of the OCP – especially those proposed in Padgett 1992 and Padgett 1995 – have begun to diverge from its traditional formulation in at least three important respects: (1) the OCP is sometimes claimed to be noncategorical, (2) formulation of this principle has in some cases shifted away from statements about featural identity and towards statements about the relative similarity of the elements in question, and (3) adjacency requirements are sometimes softened or downplayed.

All of these modifications are in the spirit of the research presented here.

7.2 Stochastic constraints

Frisch, Broe and Pierrehumbert 1995 and Frisch 1996 build on the work of Pierrehumbert 1993 to propose a novel account of the cooccurrence restrictions operating on place of articulation within Arabic verbal roots. These authors suggest that these cooccurrence restrictions are rooted in similarity; the definition of similarity to which they appeal is based on shared natural classes. In 7.2.1 and 7.2.2, I summarize the relevant aspects of this work, which I will refer to as the "stochastic constraints" analysis. I suggest that this type of analysis is incapable of handling much of the data discussed in this study: it does not appear able to account for languages in which identical elements are allowed to cooccur, although similar elements are prohibited from cooccurring. Languages of this type include Peruvian Aymara, Gojri, Bolivian Aymara, Hausa, Old Georgian, and Tzutujil.

7.2.1 Pierrehumbert 1993

Pierrehumbert 1993 addressed the cooccurrence restrictions on place of articulation in Arabic verbal roots. In that language, not all consonants freely cooccur within roots.

Arabic is well-known for its nonconcatenative morphology: roots are generally composed of three consonants, and vowels are provided by other morphemes. For example, katab 'to write' and kutib 'to be written' both come from the verbal root ktb. The vowel melodies (a and ui) and skeletal patterns (in both cases CVCVC) are supplied

by other morphemes (McCarthy 1981; see references there for relevant earlier analyses). The cooccurrence restrictions of Arabic (and of Semitic languages more generally) were described by Greenberg 1950 (see references there for earlier observations along the same lines). Greenberg pointed out that Arabic verbal roots never begin with two identical consonants, although they do sometimes end with identical consonants (*mmd, but cf. mdd 'to stretch', frr 'to flee', etc.), and that roots usually do not include homorganic consonants.

McCarthy 1979 first provided an explanation for the asymmetrical prohibition on adjacent identical consonants (i.e., for why *mmd is prohibited while mdd is allowed; this explanation is also found in McCarthy 1981, 1986). In that study, the asymmetry was analyzed as resulting from a prohibition against adjacent identical segments in the *underlying forms* of roots. This prohibition ruled out roots of type $C_1C_2C_2$ and $C_1C_1C_2$. The existence of roots that appear to be of type $C_1C_2C_2$ (for example, mdd 'to stretch') is accounted for by positing biliteral (C_1C_2) roots and left-to-right templatic association: the underlyingly biliteral roots expand to a $C_1C_2C_2$ surface pattern. Under this analysis, mdd 'to stretch' has the underlying form md. Given left-to-right association of segments to skeletal positions, of course, $C_1C_1C_2$ cannot be produced from biliteral roots.

McCarthy's explanation of the lack of Arabic verbal roots of form $C_1C_1C_2$ is widely accepted. His account of the infrequency of roots containing homorganic consonants, however, was convincingly refuted by Pierrehumbert 1993. In that article, the author examined statistical patterns in the Arabic data and detailed several ways in

which OCP-based analyses failed to capture these generalities. Under Pierrehumbert's account, the restrictions on Arabic roots are based on a dispreference for "perceived similarity": consonants become progressively less likely to cooccur as they become more similar. Perceived similarity is defined as a function of objective similarity and proximity; objective similarity is defined as in (10). The features referred to in (10) are phonological features, which Pierrehumbert assigns according to the theory of contrastive underspecification.

(10)

shared features / (shared features + unshared features)

In Arabic, only similar segments *that have the same place of articulation* are subject to cooccurrence restrictions. Therefore, similarity values are only computed for segments within classes representing the major places of articulation: labial, coronal, and dorsal/pharyngeal.

The formula given in (10) properly defines identity as the limiting case of similarity. In other words, restrictions on the cooccurrence of homorganic and identical segments are different in strength – because identical segments are more similar than homorganic segments – but not in kind.

As mentioned above, Greenberg noted that homorganic consonants, although not absolutely prohibited, tended not to cooccur within verbal roots of Arabic. It turns out that "homorganic", however, is both too strong and too weak a term to describe the

cooccurrence restrictions at hand. In Arabic, velars and uvulars (/k/g/q/χ/ʁ/) are collapsed into a single group, as are "guttural approximants" (/χ/ʁ/ħ/ʕ/h/ʔ/). Coronals, on the other hand, are divided into obstruents (/t/d/tˤ/dˤ/θ/ð/s/z/sˤ/zˤ/ʃ/) and sonorants (/l/r/n/). Coronal obstruents and coronal sonorants are subject to stronger within-group cooccurrence restrictions than are coronals as a bunch. Finally, the labial segments /b/f/m/ also tend not to cooccur.

Pierrehumbert 1993 examined these patterns and finer effects by comparing the observed and expected numbers of various consonant pairs in different positions in triliteral Arabic roots. The frequencies of occurrence of different consonants were computed on the basis of 2676 verbal roots taken from the 1979 edition of Wehr's dictionary of modern Arabic (Cowan 1979). The author found strong root-level cooccurrence restrictions for identical consonants, and weaker cooccurrence restrictions for homorganic, non-identical consonants. Distance also affected the results: adjacent consonants suffered stronger constraints against cooccurrence than did non-adjacent consonants.

Pierrehumbert 1993 was superseded by Frisch, Broe and Pierrehumbert 1995 (hereafter Frisch et al. 1995) and Frisch 1996. The authors of these studies present convincing arguments against portions of the analysis presented in Pierrehumbert 1993, in particular the use of contrastive underspecification in assigning feature values to segments, and reliance on phonological features rather than natural classes in the

computation of similarity values. These two studies are the subject of the following section.

7.2.2 Frisch et al. 1995 and Frisch 1996

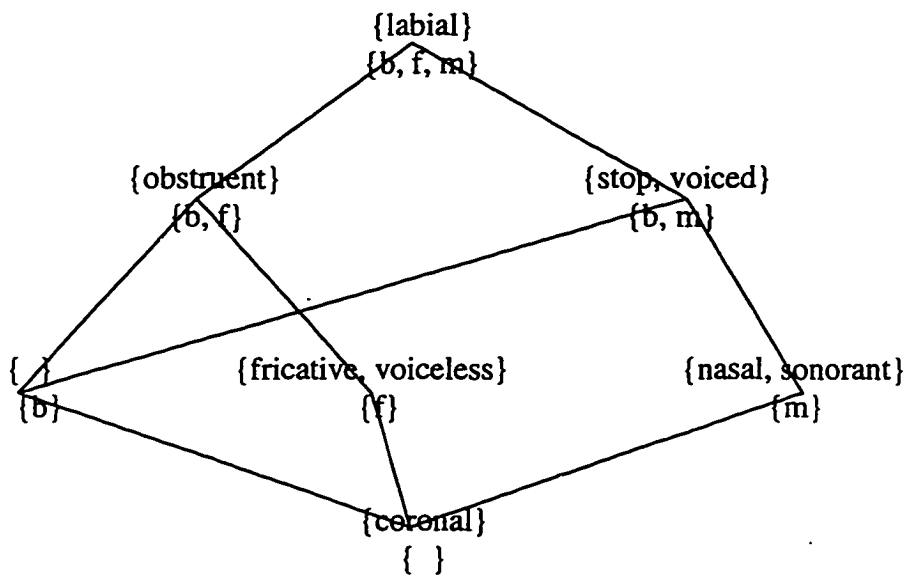
Frisch et al. 1995 and Frisch 1996 work with the data patterns established in Pierrehumbert 1993 to produce a more satisfactory definition of similarity than that provided by the formula shown in (10). In Frisch et al. 1995, the authors offer a definition of similarity based on natural classes; a "stochastic constraint" models acceptability vs. similarity with a logistic function. In this work, natural classes are constructed for the phoneme inventory on the basis of the theory of structured specification set forth in Broe 1993. In the rest of this section, I describe how natural classes and similarity are computed in these analyses. I then give an example of a stochastic constraint, and explain why such constraints cannot capture the identity effect.

Structured specification theory models the segment inventory using "tangled hierarchies" or "lattices". These lattices allow non-redundant, partially redundant, and redundant features to contribute to similarity in different degrees. Non-redundant features increase similarity to a greater degree than do partially redundant features, while redundant features do not contribute to similarity at all. An example of a lattice modeling a small set of phonemes – the labial segments /b/, /f/, and /m/ – is given in (12) (after Figure 6a in Frisch et al. 1995). The feature values represented by the lattice in (12) are given in the chart in (11).

(11)

	labial	obstruent	stop	voiced	voiceless	fricative	nasal	sonorant	coronal
b	+	+	+	+					
f	+	+			+	+			
m	+		+	+			+	+	

(12)



The lattice in (12) allows one to easily read off the sets of natural classes that characterize each of the segments contained in it. The segment /m/ falls into three natural classes: {[labial]}, {[stop], [voiced], [labial]}, and {[nasal], [sonorant], [stop], [voiced], [labial]}. The segment /f/ also falls into three natural classes: {[labial]}, {[obstruent], [labial]}, and {[fricative], [voiceless], [obstruent], [labial]}. The segment /b/ falls into four natural classes: {[labial]}, {[obstruent], [labial]}, {[stop], [voiced], [labial]}, and {[stop], [voiced], [obstruent], [labial]}.

The only natural class that is shared by /f/ and /m/ is that defined by {[labial]}; /b/ and /m/ share the two natural classes {[stop], [voiced], [labial]} and {[labial]}, while /f/ and /b/ share the two natural classes {[obstruent], [labial]} and {[labial]}.

Notice that completely redundant features do not create new natural classes. The feature [coronal] is redundant because it does not characterize any of the segments shown in the lattice in (12). Therefore, it does not create any additional natural classes.

In Arabic, cooccurrence restrictions operate over similar *homorganic* segments. Therefore, Frisch et al. 1995 and Frisch 1996 only compute similarity within sub-lattices representing each major place of articulation. In those two studies, similarity is calculated according to the equation shown in (13). A value of 0 would indicate complete dissimilarity – the pair of segments under consideration share no natural classes. A value of 1 would indicate maximal similarity – in other words, the pair of segments under consideration are identical.

(13)

shared natural classes / (shared + non-shared natural classes)

The similarity values of the three segments represented in (12) are shown in (14).

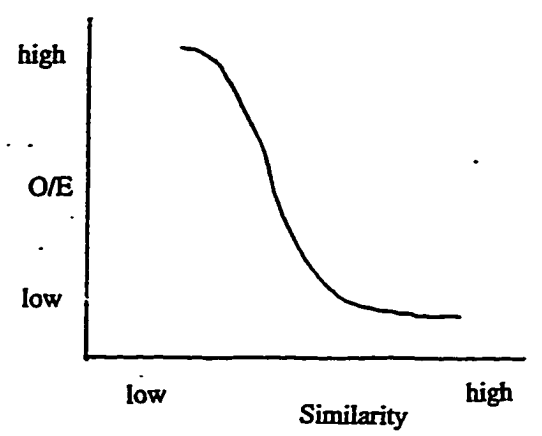
(14)

	b	f	m
b	1	0.4	0.4
f		1	0.2
m			1

Each segment is maximally similar to itself: the pair {b,b}, for example, has a similarity value of 1. The pairs {b, m} and {b, f} are equivalently similar (similarity = 0.4); the pair {f,m} is less similar (similarity = 0.2).

In Frisch et al. 1995, the actual and expected numbers of roots containing different pairs of consonants were computed, as in Pierrehumbert 1993. The ratios of observed to expected values were then calculated for roots of various types (O/E), and these values were graphed against similarity values. A low O/E indicates the presence of a cooccurrence restriction; a high O/E indicates the absence of a cooccurrence restriction. The resulting plot was fit with a logistic function resembling that shown in (15) (adapted from Figure 8, Frisch et al. 1995[3]).

(15)



[3] Frisch et al. 1995 actually provides separate functions for segments that are adjacent in the root and segments that are not adjacent in the root. Frisch 1996 provides three functions, one describing the restriction operating over segments in the first two positions of the root, one describing the restriction operating over segments in the second two positions of the root, and a third describing the restriction operating over non-adjacent segments.

The function shown in (15) represents a probabilistic, or "stochastic" constraint. The x-axis describes a continuum ranging from very dissimilar to maximally similar (i.e. identical) segments. Frisch et al. suggest that the y-axis, which represents the likelihood that a form of a given type will be found, can also be understood to represent acceptability, or wellformedness. Rather than governing the mapping of input forms to output forms, stochastic constraints describe the relative numbers of forms of different types in the lexicon. In Arabic, forms containing homorganic segments are underattested. The more similar a pair of homorganic segments are (according to the natural classes similarity metric), the less likely they are to cooccur.

The lexicon of modern Arabic matches the predictions of the stochastic constraint model fairly closely. Of course, the constraint sketched in (15) was constructed to fit the Arabic lexicon; the important result of this method is that a single method of calculating similarity – natural classes similarity, computed within sub-lattices representing the major places of articulation – can serve as the basis of constraints describing the occurrence of roots of various types within the lexicon.

The research described in Frisch et al. 1995 and Frisch 1996 differs from the analysis pursued here in several respects. First, the stochastic constraint model describes statistical characteristics of existing lexicons, and shows how they correlate with a similarity metric that is based on natural classes. The model explored here instead addresses relations between the mappings of input and output forms. (I also suggest that it is auditory similarity, rather than natural classes similarity, that lies at the base of

laryngeal cooccurrence restrictions; however, I do not have experimental evidence supporting this conjecture.)

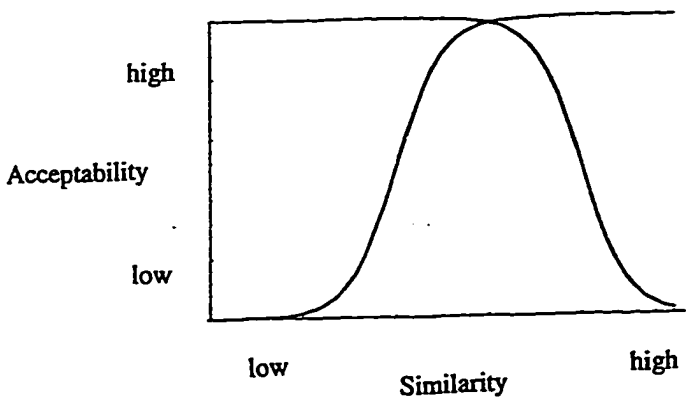
Second, the stochastic constraint model properly pays attention to gradient data patterns. The research pursued here explores categorical restrictions in a variety of languages (and also treats near-categorical restrictions as categorical); gradiency is treated only as it arises in the implicational hierarchy that governs categorical cooccurrence restrictions of different languages. The existence of gradient restrictions within the languages discussed in this study is an important topic for future research; gradient data on laryngeal cooccurrence restrictions would provide an excellent test of the natural classes-based similarity metric developed to handle the place cooccurrence restrictions of Arabic. (See Berkley 1994 for a discussion of the problems gradient data present to standard Optimality Theory analyses.)

Finally, the stochastic account makes no provision for the identity effect. In the Arabic data, identity is relevant only as the limiting case of similarity: similar segments tend not to cooccur, and identical segments – which are of course the most similar – do not cooccur at all. In many of the languages treated here, however (for example, Peruvian Aymara, Gojri, Bolivian Aymara, Hausa, Old Georgian, and Tzutujil), similar segments are disfavored from cooccurring *unless they are identical.*

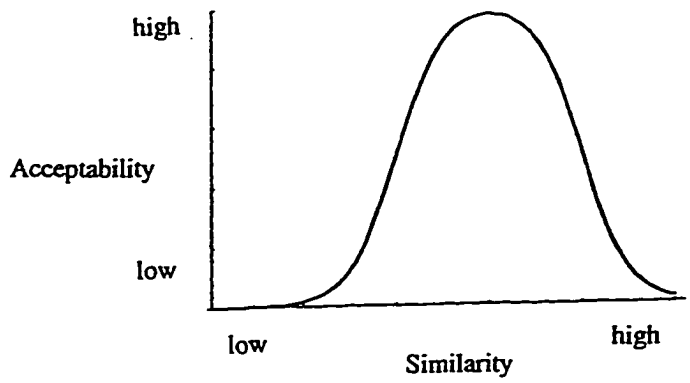
The mechanism of stochastic constraint combination is multiplication (Frisch et al. 1995, Frisch 1996). Frisch et al. offer parasitic vowel harmony as an example of a phonological effect that can be modelled through stochastic constraint combination. The

authors state that in languages with parasitic vowel harmony, roots are most likely to contain moderately similar vowels. In the analysis of Frisch et al., this results from the combination of a constraint penalizing similarity with a constraint penalizing dissimilarity. As the authors note, combining two such constraints will produce a bell-shaped curve, in which moderate degrees of similarity are the most acceptable result. An assimilitory constraint and a dissimilatory constraint are shown in (16); the product of these constraints is shown in (17) (adapted from Figures 12 and 13 in Frisch et al. 1995).

(16)

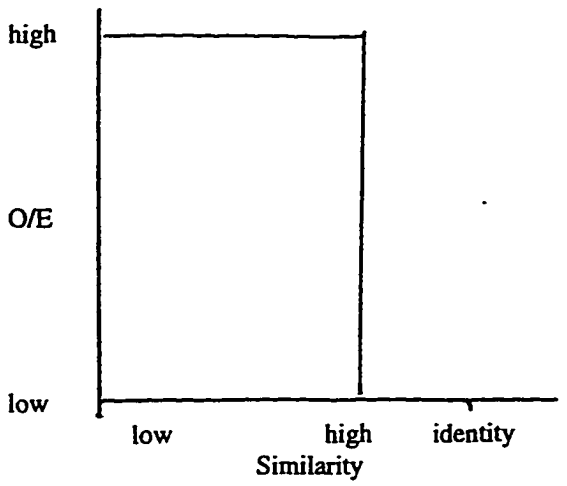


(17)



Consider, now, how one might attempt to account for the identity effect within the framework of a stochastic constraint analysis. The analysis would require a dissimilarity constraint – a constraint that penalizes similar segments. Such a constraint is shown in (18) (because I have not collected information on gradient restrictions within the languages treated here, the function shown is one that would be appropriate to modelling a categorical constraint).

(18)

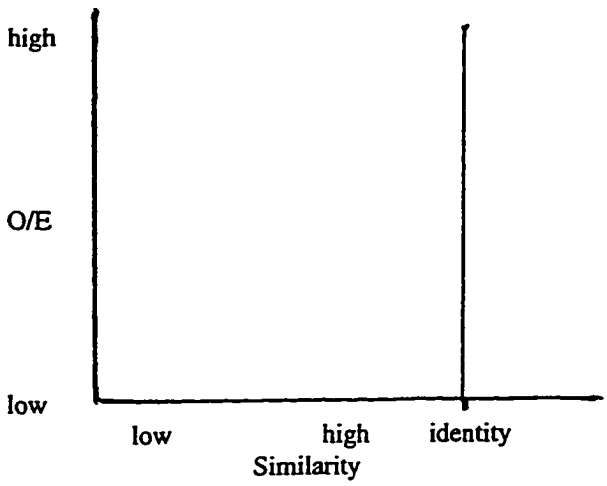


Constraint penalizing similar segments

Notice that the value of O/E drops to 0 for very similar pairs of consonants. This constraint describes a language in which dissimilar and similar segments freely cooccur, but very similar segments do not cooccur.

Languages with the identity effect do allow identical segments to coöccur. A constraint prohibiting nonidentical segments but allowing identical segments is shown in (19).

(19)



Constraint prohibiting non-identical segments

Clearly, multiplying the constraints shown in (18) and (19) will produce a situation in which no forms are acceptable. The heart of the problem is that the two mandates instantiated by the constraints – (1) segments cannot cooccur if they are very similar or identical (see (18)), and (2) segments can only cooccur if they are identical (see (19)) – can never be satisfied by the same form. I do not know how a stochastic constraint analysis could account for the identity effect.

Frisch et al. 1995 and Frisch 1996 address place cooccurrence restrictions. It is not the case that these restrictions are exempt from the identity effect. Javanese, for example, is a language in which homorganic segments are disfavored from occurring within roots, while identical segments are allowed to cooccur (Uhlenbeck 1950, Mester 1986). Analyses of place cooccurrence restrictions, then, must allow for the existence of the identity effect, just like analyses of laryngeal cooccurrence restrictions.

Frisch 1996 does address one example of a place cooccurrence restriction that seems, on the face of it, to involve the identity effect. This is the cooccurrence restriction operating over similar homorganic segments in English.

Berkley 1994 found that monomorphemic monosyllables of English containing homorganic[4] consonants separated by a short vowel (i.e., *king*, *fib*, *run*, etc.) were underrepresented in the lexicon. Monosyllables containing homorganic consonants separated by more segments were also underrepresented, although the cooccurrence restriction grows weaker as more material intervenes. Berkley did not treat identical segments separately from homorganic but non-identical segments.

Frisch 1996 points out that in fact, identical segments do occur commonly in monomorphemic monosyllables of English: consider *pip*, *bob*, *mom*, *tat*, *dad*, *nun*, *kick*, *gag*, etc. That author reports on a study using the CELEX English-language on-line dictionary. He looked at identical singleton onset and coda consonants in monosyllables, and in the initial syllables of polysyllabic words. Frisch found that, with the exception of /r/ and /l/, identical onsets and codas in stressed syllables are subject only to very weak cooccurrence restrictions (/r/ and /l/ are subject to stronger cooccurrence restrictions). In unstressed syllables, by contrast, identical segments are much less likely to cooccur.

[4] Berkley found that coronal obstruents and coronal sonorants acted separately within these restrictions, so "homorganic" actually refers to effects within four segment groupings: labial consonants, coronal obstruents, coronal sonorants, and dorsal consonants.

The explanation suggested by Frisch is that stress enhances positionally-dependent allophonic variation, rendering onset and coda allophones distinct enough that they are not subject to the similarity-based cooccurrence restriction on segments at identical places of articulation. The problem with this account (which strays from the natural classes-based measure of similarity proposed for Arabic) is that homorganic but non-identical segments in stressed syllables *are* subject to cooccurrence restrictions. This analysis would suggest that stressed-syllable onset [f] and coda [b], for example (as in *fib*), are more similar than stressed-syllable onset [b] and coda [b] (as in *bob*). This seems unlikely.

The stochastic constraints analysis provides the best model yet proposed for the place cooccurrence restrictions of Arabic. However, in order to provide a complete account of cooccurrence restrictions, the theory would need to be extended to cover the identity effect.

8: CONCLUSION

This research has addressed the existence of laryngeal cooccurrence restrictions in several languages: Cuzco Quechua, Souletin Basque, Sanskrit, Peruvian Aymara, Ofo, Gojri, Bolivian Aymara, Hausa, Old Georgian, Tzutujil, and Shuswap. I have shown that traditional OCP-based accounts, although successful in analyzing data from some languages, cannot be extended to cover everything in the typological statement.

I have suggested instead that laryngeal cooccurrence restrictions are based on auditory similarity, and I have shown that they fall into four distinct patterns. These patterns fit an implicational hierarchy: the Pattern One restrictions are a subset of the Pattern Two restrictions, which are in turn a subset of the Pattern Three restrictions, and so on.

The data also shows the existence of two types of identity effect. The complete version of the identity effect is found in Peruvian Aymara, Gojri, Bolivian Aymara, Hausa, Old Georgian, and Tzutujil. In these languages, all identical elements (from the set of elements over which laryngeal cooccurrence restrictions operate) may cooccur. For example, hahi, $t^h at^h u$, and k'ik'a are all possible forms in both Aymara dialects. The partial version of the identity effect is manifested in Shuswap and, I have argued here, in Ofo as well. In languages with the partial identity effect, two /h/s can cooccur within morphemes, as may two glottal stops, but identical aspirated or ejective stops cannot cooccur.

I have accounted for these facts with an Optimality Theoretic analysis. In this analysis, the phenomena of laryngeal cooccurrence restrictions emerge from the interactions of a constraint that penalizes auditory similarity, a constraint that requires identity, and constraints that enforce the correspondence of laryngeal features between input and output forms.

REFERENCES

Abraham, R.C. 1962 [1946]. *Dictionary of the Hausa language*. Hodder and Stoughton.

Adelaar, Willem. 1986. Relacion quechua-aru. *Revista Andina* 4:379-425.

Allen, W.S. 1957. Aspiration in the Harauti nominal. *Studies in linguistic analysis, Philological Society*, 68-86. Oxford: Blackwell.

Andersen, Torben. 1993. Aspects of Berta phonology. *Afrika und Ubersee* 76:41-80.

Anderson, Stephen R. 1970. On Grassmann's Law in Sanskrit. *Linguistic Inquiry* 1:387-396.

Ayala Loayza, Juan Luis. 1988. *Diccionario espanol-aymara aymara-espanol*. Consejo Nacional de Ciencia y Tecnologia.

Beasley, David, and Kenneth L. Pike. 1957. Notes on Huambisa phonemics. *Lingua Posnaniensis* 6:1-8.

Beeler, M.S. 1970. Sibilant harmony in Chumash. *International Journal of American Linguistics* 36:14-17.

Berkley, Deborah Milam. 1994. Variability in Obligatory Contour Principle effects. *Papers from the 30th Regional Meeting of the Chicago Linguistic Society vol. 2: the Parasession on Variation in Linguistic Theory*, 1-12. Chicago Linguistic Society: Chicago, Illinois.

Borowsky, Toni, and Ralf-Armin Mester. 1983. Aspiration to roots remarks on the Sanskrit diaspirates. *CLS 19. Papers from the Nineteenth Regional Meeting*, 52-63.

Briggs, Lucy Therina. 1985. Dialectal variation in Aymara. *South American Indian Languages*, edited by Harriet E. Manelis Klein and Louisa R. Stark, 595-616. University of Texas Press.

Broe, Michael. 1993. *Specification theory*. PhD dissertation, University of Edinburgh.

Campbell, Lyle. 1973. On glottalic consonants. *International Journal of American Linguistics* 39:44-46.

Carenko, E.I. 1975 [1972]. On laryngealization in Quechua. *Linguistics* 146:5-14.

Cerulli, E. 1947. Three Berta dialects in western Ethiopia. *Africa*. (Reprinted in Tucker and Bryan 1966).

Cherkesi, E. 1950. *Georgian-English dictionary*. Oxford.

Cowan, J. 1979. *Hans Wehr: a dictionary of modern written Arabic*. Wiesbaden: Otto Harrasowitz.

Crapo, Richley H., and Percy Aitken. 1986. *Bolivian Quechua reader and grammar-dictionary*. Ann Arbor: Karoma Publishers, Inc.

Cusihuaman, Antonio G. 1976a. *Gramatica quechua: cuzco-collao*. Lima, Peru: Ministerio de Educacion.

Cusihuaman, Antonio. 1976b. *Diccionario quechua: cuzco-collao*. Ministerio de Educacion Instituto de Estudios Peruanos.

Dayley, J. 1985. *Tzutujil grammar*. University of California Press.

De Lucca, Manuel. 1987. *Diccionario aymara-castellano castellano-aymara*. La Paz, Cochabamba: Editorial Los Amigos del Libro.

De Reuse, Willem J. 1981. Grassmann's Law in Ofo. *International Journal of American Linguistics* 47:243-244.

Deza Galindo, Juan Francisco. 1989. *Diccionario aymara-castellano castellano-aymara*. Consejo Nacional de Ciencas y Tecnologica.

Dorsey, James O., and John R. Swanton. 1912. *A dictionary of the Biloxi and Ofo languages*. Washington Government Printing Office, U.S. Bureau of American Ethnology Bulletin #47.

Emmorey, Karen Denise. 1987. *Morphological structure and parsing in the lexicon*. PhD dissertation, UCLA.

Fahnrich, Heinz. 1991. Old Georgian. *The indigenous languages of the Caucasus, vol. 1: the Kartvelian languages*, edited by Alice C. Harris, 131-217. Caravan Books.

Flemming, Edward. 1995. *Auditory representations in phonology*. PhD dissertation, UCLA.

Fordyce, James F. 1980. On the nature of glottalic and laryngealized consonant and vowel systems. *UCLA Working Papers in Phonetics* 50:120-154.

Frisch, Stefan. 1996. *Similarity and frequency in phonology.* PhD dissertation, Northwestern University.

Frisch, Stefan, Michael Broe, and Janet Pierrehumbert. 1995. *The role of similarity in phonotactic constraints.* Unpublished manuscript, Northwestern University.

Goldsmith, John. 1976. *Autosegmental phonology.* Indiana University Linguistics Club.

Grassmann, Hermann. 1863. Uber die Aspiraten und ihr gleichzeitiges Vorhandensein in An- und Auslaute der Wurzeln. *Seitschrift fur vergleichende Sprachforschung auf dem Gebiete des Deutschen, Griechischen und Lateinischen* 12:81-138.

Greenberg, Joseph H. 1950. The patterning of root morphemes in Semitic. *Word* 6:162-182.

Greenberg, Joseph H. 1970. Some generalizations concerning glottalic consonants, especially implosives. *International Journal of American Linguistics* 36:123-145.

Grimes, Barbara F., editor. 1988. *Ethnologue*, eleventh edition. Dallas, Texas: Summer Institute of Linguistics, Inc.

Haas, Mary R. 1969. Swanton and the Biloxi and Ofo dictionaries. *International Journal of American Linguistics* 35:286-290.

Hardman, M.J. 1985. Aymara and Quechua: languages in contact. *South American Indian languages*, edited by Harriet E. Manelis Klein and Louisa R. Stark, 617-643. University of Texas Press.

Hardman, M.J., J. Vasquez, and J. de D. Yapita Moya, editors. 1974. *The Aymara Language Project, volume III: outline of Aymara phonological and grammatical structure.* Gainesville: Department of Anthropology, University of Florida.

Haudricourt, A.G. 1950. Les consonnes preglottalisees en indochine. *Bulletin de la Societe de Linguistique de Paris* 46:172-182.

Hoard, James. 1975. On incorporating Grassmann's Law into Sanskrit Phonology. *Proceedings of the first annual meeting of the Berkeley Linguistics Society*, edited by Cathy Cogen, Henry Thompson, Graham Thurgood, Kenneth Whistler, and James Wright, 207-220. Berkeley, California: University of California.

Hoijer, Harry and Edward P. Dozier. 1949. The phonemes of Tewa, Santa Clara dialect. *International Journal of American Linguistics* 15:139-144.

Hopper, Paul. 1973. Glottalized and murmured occlusives in Indo-European. *Glossa* 7:141-166.

Hornberger, Esteban S. and Nancy H. Hornberger. 1983. *Diccionario tri-lingüe Quechua de Cusco: Quechua English Castellano.* La Paz, Bolivia (xerox).

Hualde, Jose Ignacio. 1993. Topics in Souletin phonology. *Generative studies in Basque linguistics*, Jose Ignacio Hualde and Jon Ortiz de Urbina, 289-327. John Benjamins Publishing Company.

Ito, Junko, and Ralf-Armin Mester. 1986. The phonology of voicing in Japanese: theoretical consequences for morphological accessibility. *Linguistic Inquiry* 17:49-73.

Javkin, Hector. 1977. Towards a phonetic explanation for universal preferences in implosives and ejectives. *Proceedings of the 3rd Annual Meeting of the Berkeley Linguistics Society*, 559-565.

Kaufman, Terence. 1974. *Idiomas de Mesoamerica.* Guatemala: Seminario de Integracion Social Gualtemateca Publicacion.

Kaun, Abigail. 1995. *The typology of rounding harmony: an optimality theoretic approach.* PhD dissertation, UCLA.

Kawasaki, Haruko. 1982. *An acoustical basis for universal constraints on sound sequences.* PhD dissertation, University of California, Berkeley.

Kingston, John. 1985. *The phonetics and phonology of the timing of oral and glottal events.* PhD dissertation, UC Berkeley.

Kiparsky, Paul. 1965. *Phonological change.* PhD dissertation, MIT.

Kraft, Charles H. and Marguerite G. Kraft. 1973. *Introductory Hausa.* University of California Press.

Kuipers, Aert H. 1974. *The Shuswap language*. Paris: Mouton.

Ladefoged, Peter L. 1964. *Phonetic differences in West African languages*. Cambridge: Cambridge University Press.

Lafon, Rene. 1958. Contribution a l'etude phonologique du parler basque de Larrau (Haute-Soule). *Estructuralismo e historia: miscelanea homenaje a Andre Martinet*, vol. 2., edited by Diego Catalan, 77-106. La Laguna, Canarias: Universite de La Laguna.

Landerman, Peter. 1994. Glottalization and aspiration in Quechua and Aymara reconsidered. *Language in the Andes*, edited by Peter Cole, Gabriella Hermon and Mario Daniel Martin, 332-378.

Larrasquet, Jean. 1939. *Le basque de la Basse-Soule orientale*. Paris: La Societe de Linguistique de Paris.

Leben, William. 1973. *Suprasegmental phonology*. PhD dissertation, MIT.

Lehmann, Winifred P. 1966. *Proto-Indo-European phonology*. Austin, Texas: The University of Texas Press and Linguistic Society of America.

Lombardi, Linda. 1991. *Laryngeal features and laryngeal neutralization*. PhD dissertation, University of Massachusetts.

Lyman, Benjamin Smith. 1894 [1888]. The change from surd to sonant in Japanese compounds. *A selection of the papers read before The Oriental Club of Philadelphia*, 160-176. Boston: Ginn & Company.

MacEachern, Margaret R. 1996. Ordering restrictions on aspirated and ejective stops in Aymara. *Proceedings of the Twenty-Second Annual Meeting of the Berkeley Linguistics Society*, edited by Jan Johnson, Matthew L. Juge, and Jeri L. Moxley, 244-254. Berkeley, California: Berkeley Linguistics Society.

MacEachern, Margaret R. 1997. Laryngeal similarity effects in Quechua and Aymara. *The Proceedings of the Fifteenth West Coast Conference on Formal Linguistics*, edited by Brian Agbayani and Sze-Wing Tang, 337-351. Center for the Study of Language and Information, Stanford University.

Maddieson, Ian. 1984. *Patterns of Sounds*. Cambridge University Press.

Mannheim, Bruce. 1991. *The language of the Inka since the European invasion*. Austin: The University of Texas Press.

Marslen-Wilson, W. 1984. Function and process in spoken word recognition: A tutorial review. *Attention and performance X: Control of language processes*, edited by H. Bouma and D. Bouwhuis. London: Lawrence Erlbaum Associates.

Marslen-Wilson, W. 1987. Functional parallelism in spoken word recognition. *Cognition* 25:71-102.

Marslen-Wilson, W., and L.W. Tyler. 1980. The temporal structure of spoken language understanding. *Cognition* 8:1-71.

Marslen-Wilson, W., and A. Welsh 1978. Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology* 10:29-63.

Martin, Laura. 1988. Fonologia. *Aymara compendio de estructura fonologica y gramatical*, by M.J. Hardman, Juana Vasquez and Juan de Dios Yapita, 24-66. La Paz, Bolivia: Instituto de Lengua y Cultura Aymara.

Matthews, G.H. 1958. *Handbook of Siouan languages*. PhD dissertation, University of Pennsylvania.

McCarthy, John. 1979. *Formal problems in Semitic phonology and morphology*. PhD dissertation, MIT.

McCarthy, John. 1981. A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry* 12:373-418.

McCarthy, John. 1986. OCP effects: gemination and antigemination. *Linguistic Inquiry* 17:207-263.

McCarthy, John. 1988. Feature Geometry and dependency: a review. *Phonetica* 43:84-108.

McCarthy, John. 1989. Linear order in phonological representation. *Linguistic Inquiry* 20:71-99.

McCarthy, John, and Alan Prince. 1993. *Prosodic morphology I: constraint interaction and satisfaction*. Unpublished manuscript, University of Massachusetts, Amherst, and Rutgers University.

McCarthy, John, and Alan Prince. 1994. The emergence of the unmarked: Optimality in Prosodic Morphology. *Proceedings of the North East Linguistics Society* 24, edited by Merce Gonzalez, 333-379. GLSA, University of Massachusetts at Amherst.

McCarthy, John, and Alan Prince. 1995. Faithfulness and reduplicative identity. *University of Massachusetts Occasional Papers 18: Papers in Optimality Theory*, edited by Jill N. Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk, 249-384. GLSA, University of Massachusetts at Amherst.

Mester, Armin. 1986. *Studies in tier structure*. PhD dissertation, University of California, San Diego.

Michelena, Luis. 1976. *Fonetica historica Vasca*, second edition. San Sebastian: Publicaciones del Seminario Julio de Urquijo de la Excma. Diputacion Provincial de Guipuzcoa.

Miranda, R.P. Pedro. 1970. *Diccionario breve castellano-aymara y aymara-castellano*. La Paz, Bolivia.

Molitor, Joseph. 1952. *Altgeorgisches Glossar zu ausgewahlten Bibeltexten*. Rome, Pontificium Institutum Biblicum.

Ohala, John, and C.J. Riordan. 1979. Passive vocal tract enlargement during voiced stops. *Speech Communication Papers from the 97th Meeting of the Acoustical Society of America*, edited by J.J. Wolf and D. Klatt.

Orr, Carolyn, and Robert E. Longacre. 1968. Proto-Quechumaran. *Language* 44:528-554.

Padgett, Jaye. 1992. OCP subsidiary features. *Proceedings of the North East Linguistic Society* 22, 335-346. University of Massachusetts.

Padgett, Jaye. 1995. *Stricture in Feature Geometry*. Stanford, California: CSLI Publications.

Parker, Steve, and David Weber. 1996. Glottalized and aspirated stops in Cuzco Quechua. *International Journal of American Linguistics* 62:70-85.

Parsons, F.W. 1970. Is Hausa really a Chadic language? Some problems of comparative phonology. *African Language Studies* XI:272-288. School of Oriental and African Studies, University of London.

Pater, Joe. Forthcoming. Austronesian nasal substitution and other NC̥ effects. *The prosody morphology interface*, edited by Rene Kager, Harry van der Hulst, and Wim Zonneveld. Cambridge University Press.

Phelps, E. 1975. Sanskrit diaspirates. *Linguistic Inquiry* 6:447-464.

Phelps, E., and M. Brame. 1973. On local ordering of rules in Sanskrit. *Linguistic Inquiry* 4:387-400.

Pierrehumbert, Janet. 1993. Dissimilarity in the Arabic verbal roots. *Proceedings of the North East Linguistic Society* 23, 368-381. University of Ottawa.

Prince, Alan, and Paul Smolensky. 1993. *Optimality Theory: constraint interaction in generative grammar*. Unpublished manuscript, Rutgers University and University of Colorado, Boulder.

Reidhead, P.W. 1946. Report of linguistic survey among tribes Berta, Ingassana, Koma, Uduk, Jum Jum, Maban. (Reprinted in Tucker and Bryan 1966.)

Rood, David S. 1979. Siouan. *The languages of Native America*, edited by Lyle Campbell and Marianne Mithun, 236-298. University of Texas Press.

Rowe, J.H. 1950. Sound patterns in three Inca dialects. *International Journal of American Linguistics* 16:137-148.

Ruhlen, Merritt. 1975, 1976. *A guide to the languages of the world*. Stanford Language Universals Project, Stanford University.

Sag, Ivan. 1974. The Grassmann's Law Ordering Pseudoparadox. *Linguistic Inquiry* 5:591-607.

Sag, Ivan. 1976. Pseudosolutions to the pseudoparadox: Sanskrit diaspirates revisited. *Linguistic Inquiry* 7:609-622.

Schindler, Jochem. 1976. Diachronic and synchronic remarks on Bartholomae's and Grassmann's Laws. *Linguistic Inquiry* 7:622-637.

Sharma, Jagdish Chander. 1979. *Gojri phonetic reader*. Phonetic Reader Series 19. Mysore, India: Central Institute of Indian Languages.

Shaw, Patricia A. 1991. Consonant harmony systems: the special status of coronal harmony. *Phonetics and phonology* volume 2, 125-157.

Smolensky, Paul. 1996. The initial state and 'richness of the base' in Optimality Theory. *Johns Hopkins University Technical Report* JHU-CogSci-96-4.

Swanton, John R. 1909. A new Siouan dialect. *Putnam Anniversary Volume Anthropological Essays*, 477-486.

Thompson, Laurence C., and M. Terry Thompson. 1985. A Grassmann's Law for Salish. *For Gordon Fairbanks*, edited by Veneeta Z. Acson and Richard L. Leed, 143-147. University of Hawaii Press, OL, Special Publication #20.

Torero Fernandez de Cordova, Alfredo. 1964. Los dialectos quechuas. *Anales Cientificos de la Universidad Agraria* 2:446-478.

Triulzi, A., A.A. Dafallah, and M. Lionel Bender. 1976. Berta. *The non-Semitic languages of Ethiopia*, edited by M. Lionel Bender, 513-532. Monograph No. 5, Occasional Papers Series, Committee on Ethiopian Studies. African Studies Center, Michigan State University.

Trubetzkoy, N.S. 1969 [1958]. *Principles of phonology*. Translated by Christiane A.M. Baltaxe. University of California Press.

Tucker, A.N., and Margaret A. Bryan. 1966. *Linguistic analyses: the non-Bantu languages of northeastern Africa*. London: OUP for IAI.

Uhlenbeck, E.M. 1950. The structure of the Javanese morpheme. *Lingua* 2:239-270.

Vogt, Hans. 1988. Structure phonemique du georgien. *Studia Caucasologica II*, edited by Even Hovdhaugen and Fridrik Thordarson, 344-429. Norwegian University Press.

Wang, William S.Y. 1968. *The basis of speech (draft copy)*. Project on Linguistic Analysis, Phonology Laboratory, Department of Linguistics, University of California, Berkeley. Second Series, No. 4.

Warkentin, Viola M., and Ruth M. Brend. 1974. Chol phonology. *Linguistics* 132:87-101.

Weathers, Nadine. 1947. Tsotsil phonemes with special reference to allophones of b. *International Journal of American Linguistics* 13:108-111.

Whitney, William Dwight. 1941 [1889]. *Sanskrit grammar*. Cambridge, Massachusetts: Harvard University Press.

Whitney, William Dwight. 1945 [1885]. *The roots, verb-forms, and primary derivatives of the Sanskrit language*. New Haven, Connecticut: American Oriental Society.

Wolff, Hans. 1950. Comparative Siouan I. *International Journal of American Linguistics* 16:61-66.

Wolff, Hans. 1952. Osage K: phonemes and historical phonology. *International Journal of American Linguistics* 18:63-68.

Yip, Moira. 1988. The Obligatory Contour Principle and phonological rules: a loss of identity. *Linguistic Inquiry* 19:65-100.

Yip, Moira. 1989. Feature Geometry and cooccurrence restrictions. *Phonology* 6:349-375.

Zue, V.W. 1980. *Acoustic characteristics of stop consonants: a controlled study.* Bloomington: Indiana University Linguistics Club.