PROSODIC PHONOTACTICS

A Dissertation Presented

by

TIM D. SHERER

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May, 1994

Department of Linguistics

Approved as to style and content by:

John J. McCarthy, Chair

Shmuel Bolozky, Member

F. Roger Higgins, Member

Elisabeth O. Selkirk, Member

PREFACE TO THE COMPUTER-FILE VERSION

This version of my dissertation is divided into seven chapters, numbered 0 to 6. Chapters 1 through 6 contain the corresponding chapters of my dissertation. Chapter 0 contains the front matter (acknowledgements, abstract, and so forth), as well as the appendix and the sources cited.

This version of my dissertation was taken from the Graduate Linguistics Student Association (GLSA) version, which had several typos corrected over the deposit version. This version differs from the GLSA version with respect to some diagrams. Mainly in tableaux, I have substituted bracketing and subscripted morae for tree diagrams in representing syllables, as below:

$\{C\ V_{\mu}\}\ \{C\ V_{\mu}\}$	=	2 cv syllables
$\{C\ V_{\mu\mu}\}$	=	a syllable with a long vowel
$\{C\ V_{\mu}\ C_{\mu}\}$	=	a bi-moraic closed syllable
$\{C\ V_{\mu}\ C\}$	=	a mono-moraic closed syllable
$\{C\ V_{\mu}\ \{\ C_{\mu}\ \}\ V_{\mu}\ \}$	=	2 syllables with a geminate consonant

There are still some diagrams in their original forms. I hope that these are recoverable from context.

The pagination is different from the original version. I have chosen to paginate the chapters separately. There is a table of contents for each chapter at the beginning of that chapter.

The first chapter summarizes the goals of the dissertation and introduces most of the constraints used later in the work. The main goal of this work is to characterize sequences of consonants by characterizing codas in the most general way possible. Since this one of the earlier works employing Prince and Smolensky's 1993 Optimality manuscript, I include an introduction to the theory in this chapter.

The second chapter includes a factorial typology involving several of the constraints introduced in chapter 1. This typology builds on Prince and Smolensky's factorial typology by employing a more detailed syllable structure. Also, there are analyses of facts from the languages Wiyot and Italian. The constraint *LONG-VOWEL is introduced in this chapter (and plays a role in chapter 4) Another question dealt with here is the content of underlying form--in particular, the result of specifying a "weight by position" mora in the lexicon. I also discuss the role of lexical levels in OT.

The third chapter deals with intervocalic consonant sequences of Finnish. A search of a machine readable dictionary revealed a coronal asymmetry in Finnish. Essentially, in Finnish a hypercharacterized syllable may be closed by a geminate or by a coronal consonant, but not by a non-coronal, non-geminate. An Optimality Theoretic analysis of this fact is offered. One aspect of this analysis is that no root not meeting the coronal only requirement would be posited by a Finnish language learner (Stampean Occultation).

The fourth chapter includes analyses of two languages, Fula and Koya, both of which permit long vowels before sequences of consonants, but not before geminates. This is analyzed by means of a syllable appendix--a consonant directly adjoined to the syllable node without an intervening mora. This version of an appendix consonant differs from the word-final version. Both Fula and Koya have cases where we might expect long vowels to precede geminates. In Fula, these cases surface with a simple consonant (corresponding to the geminate), while in Koya, these forms surface with a short vowel. Optimality Theoretic analyses for these mappings are given.

The fifth chapter focuses on the topic of syllable appendixes and appendixal consonants. The main part of the chapter is a discussion of several (closely related) Australian languages and why they should be analyzed employing appendix consonants. I provide a sketchy Optimality Theoretic analysis. The balance of the chapter is devoted to questions about the appendix position. The idea of analyzing appendixes by means of Alignment is discussed, including how this could derive different sites for adjunction (word, syllable, and so forth.) The nature of the coronal asymmetry is discussed and a phonetic scale is offered as a possible explanation.

The sixth chapter deals with topics not dealt with earlier in the dissertation, including a comparison between the Optimality Theoretic approach and an approach where constraints are set to *on* or *off*. Also discussed is the Linked Gemination approach to the representation of geminates, which is at odds with the moraic theory of geminates. Another matter covered was the role of linking, both in connection with Linked Gemination and the Coda Condition Approach, including discussion of an approach to hypercharacterization that employs coda conditions.

Tim D. Sherer
December, 1994
TDSherer@Delphi.Com

ACKNOWLEDGMENTS

There are a number of people to whom I owe a debt of gratitude for any number of things over the last nine years, since I started taking Linguistics courses at the University of Washington. This list may well fall short. In fact, it does fall short. I can only apologize to those omitted.

John McCarthy has been an excellent advisor and thesis supervisor. I am proud to have been his student.

Lisa Selkirk was also an excellent advisor and this thesis has been much improved thanks to her feedback.

Roger Higgins' contribution to this work was depth. Beyond the obvious comments and suggestions that he offered, this work has benefitted in subtle, almost mystical, fashion from his input.

Thanks to Shmuel Bolozky for his careful reading and helpful suggestions.

In addition to these committee members, a number of faculty members at the University of Massachusetts have helped or challenged me along the way, including (but not limited to) Ellen Woolford (who saw me through my second generals paper), Francois Dell, and Emmon Bach. A full enumeration of the ways in which other faculty members have seen me through along the way is beyond the scope of these acknowledgements.

Special thanks to Kathy and Lynne for putting up with me for all this time. And for a great deal of help getting through the whole experience.

At the University of Washington, Ellen Kaisse is an excellent teacher and I well appreciate her efforts in my case. Also at the University of Washington, I had the pleasure of taking classes from Joe Emonds and Frederick Newmeyer, among others.

I would also like to thank Ellen Broselow, Mark Aronoff, Marie Huffman, and the whole gang at the State University of New York at Stony Brook, for making my first "real job" a pleasure.

Among my fellow students, special note should go to my class, Juli Carter, Rose-Marie Dechaine, Jingqi Fu, Lisa Green, Noriko Kawasaki, Paul Portner, and Sylvie Robideaux. Let's get together for dinner sometime, like the old days.

Greg Lamontagne has been a steady friend for many years. He is a fine phonologist and has a fine mind in general. I never tire of arguing phonology with him and don't think I ever will. In addition, I have learned a great deal about teaching from him. And, finally, Greg Lamontagne is Greg Lamontagne.

Thanks also to Joanna Gonsalves for extensive discussion on Psychology, which she understands and I do not.

I would also like to thank Tami Kaplan.

From my University of Washington days, I want to thank my friends Kathleen Pool, Peter Goodfellow, Jennifer Hodgdon, and Melanie Orgill, and from the days before that, Frank Mendel and Barb Ericsen. And many others.

I also owe a debt of gratitude to my family, starting with my mother, Naomi Sherer, who, in addition to being a constant source of support, did a bang-up technical editing job on my first generals paper.

My brother Mike tells one hell of a story and I am always glad to hear them. My sister Nancy has ever been a buoy to my spirits. Thanks to my brother Tom for helping me get established in the northeast and much help thereafter. Thanks to Dan for being a heck of a good brother and putting up with me as a houseguest (a lot), not to mention a extensive help in getting this dissertation printed up. And thanks to June, Jerry, Suzie, and Debbie, just for being them.

My father was expansively proud of me for going to college, although he had some initial doubts about this Linguistics stuff. He was enormously proud of me for getting into graduate school. He was also dismayed that the completion of my degree would take so long, that is, no less than five years. He was afraid, I am sure, that his health would not see him through to the end of it and, sadly, it did not. It is to his memory that this thesis is dedicated.

ABSTRACT

PROSODIC PHONOTACTICS

May, 1994

TIM D. SHERER

B.A. UNIVERSITY OF WASHINGTON

Ph.D. UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by John J. McCarthy

This dissertation deals with the possibility of having syllable final consonants, including geminate consonants, and the interaction of such consonants with syllable weight and vowel length, by means of the moraic tier. The goal of this work is to have the range of patterns of consonant occurrence and vowel length follow from simple patterns in simple interactions.

A chapter is devoted to making predictions about the phonological patterns possible in the system. These are made through the interaction of simple phonological constraints, with the adoption of Optimality Theory. Since the constraints interact in different ways in different languages, they need not be complicated themselves. Still,
they produce the full range of expected patterns. Of special importance is the fact that intersyllabic consonant clusters and geminates are predicted to cross-classify.

Two chapters are devoted to more detail analyses of the languages Finnish, Fula, and Koya. In particular, these languages differ on which types of over-filled or *hypercharacterized* syllables they permit. Because hypercharacterized syllables are more than full, they are situations where we can gage the limits of syllabic constituency. An adequate approach to syllable theory must predict not only the possibility of the syllable types found in these languages, but the variation in types found between languages.

A chapter is devoted to appendix consonants, which are non-moraic non-onset consonants. Here it is proposed that only appendix consonants display the coronal asymmetries that have been described for codas in a variety of languages. Furthermore, discussion of the structural analysis of the appendix consonants is undertaken.

The aforementioned phenomena are all underlying. Another matter taken up is the nature of gemination of consonants conditioned by prosodic factors and geminates which arise as a result of concatenation. The language Wiyot, which has no distinctive length, is analyzed using the same basic schema as the languages with distinctive consonant or vowel length. Gemination by concatenation is analyzed by means of Alignment constraints.

Finally, this dissertation includes discussion of alternative representations for consonant sequences and geminates.

TABLE OF CONTENTS

APPENDIX

THE FINNISH DATA

I chose Austerlitz's glossary for two reasons. First, it included morpheme boundaries, permitting much better control for selecting monomorphemic sequences. Second, as a glossary to a reader, I hoped that it would not include extrodinarily obscure or rare words. Paging through the glosses, this seems to be correct.

A few obvious abbreviations were left out, but loan words were included, as were proper names. Capitalization for proper names was omitted, because I used caps for certain diacritic features. This has the side effect of making it impossible to distinguish proper nouns from common nouns. Had caps been included, however, the problem would not have disappeared. For instance, compare affrika and affrican, where the first is capitalized and the second is not.

Austerlitz uses three types of morpheme boundaries for derivation, inflection and compounding. These are all treated alike for my purposes. Some main entries in Austerlitz's glossary were written without morpheme boundaries, but had morpheme boundaries included in the entry in parentheses. These main entries were sometimes different in phonetic form from the versions with boundaries. The form with morpheme boundaries was used.

The total number of examples was 9654.

A cut was performed before the first morpheme boundary. The lines were sorted and uniqued. The total line count was 3304. This is the number of roots in the sample, except that some roots will show up more than once on this, as discussed below.

A "+" sign was substituted for each vowel symbol. At this point, any short vowel will be denoted with a single "+" and any long vowel or diphthong with "++". Then any sequence of 3 +'s were changed to 2 +'s, using the Word Perfect translate utility. These cases of three vowels in a row are actually bi-syllabic, according to all standard grammars. Their inclusion could have disturbed results, in principle. There were a few, but not an extraordinary number.

Sequences of two "+"'s were translated into "+=", again using the Word Perfect translate utility. Every sequence of two vowels was now indicated with one "+" and "=" indicating length. Every short vowel was indicated with a single "+".

All monosyllables, now indicated as forms with a single "+" or with a single "+=", were discarded by grep search. The line count of non-monosyllabic roots was 2755.

A cut was made between first and second "+". This kept all consonants between the first vowels as well the "=" in those cases with a long vowel, indicating length of the preceding vowel.

The result of the cut--the consonant or consonants between the first two vowels, with length diacritic--was sorted and uniqued with the count option, creating the lists below. The lists have been reordered by hand and organized into the charts in the chapter.

Examples that were subject to vowel assimilation had the surface form of the vowel in the assimilated morpheme. This caused problems because "yhte-" and "yhtä-" would appear as separate items in a root list. This will give rise to a slight inflation in numbers. I have checked over examples after long vowels to see that the effect was not profound.

raw lists and full charts

"=" indicates that the cluster follows a long vowel

11	ltt
1	ggl
1	ttrb
1	kst
1	lpp
2	nfl
6	rtt
3	nst
3	nss
1	nsl
2	rpp
2	nsk
14	rkk
1	bstr
1	mpp
2	ntt
1	ngr
2	lst
1	mpl
1	nnkk
1	kkl
1	rrl
1	lppl
1	rst
1	ngst
1	rdv
4	nkk
1	llgr
1	llm

1	rsk
4	lkk
1	rss
1	stj
4	ngl
1	skt

100	v	38	=v
3	b	1	=b
76	t	87	=t
71	s	39	=s
52	n	81	=n
69	p	26	=p
102	r	82	=r
68	h	21	=h
9	g	3	=g
24	j	10	=j
123	l	39	=l
102	k	64	=k
63	m	50	=m
2	f		
48	d	49	=d

1	ft		
30	hd	1	=hd
1	bl		
1	dh		

1	br		
1	ld		
1	kt		
1	dj		
2	fr		
9	nh		
9	hv		
41	ht	6	=ht
13	hk	3	=hk
17	ks	3	=ks
3	kr	1	=kr
1	bv		
9	hj		
1	kp		
3	bs		
1	hl		
20	lm	2	=lm
4	lh		
17	lj		
8	lp		
25	lv		
41	nt	13	=nt
3	ls		
16	lt	5	=lt
29	lk		
2	mb		
4	nj		
21	mp		
6	nd		
17	nk	1	=nk
15	ns	1	=ns
1	nm		
2	nr		
15	ng		
0	th	1	=th
2	tj		
2	tl		
25	tk		
0	tc	1	=tc
1	tm		
13	hm		
1	tv		
2	tn		
32	rv		
50	rk		

4	ps		
1	rn		
22	rm	3	=rm
1	pl		
1	pt		
22	rh		
32	rj		
2	rg		
2	tr	1	=tr
8	ts	9	=ts
1	sn		
3	sp		
2	sv		
3	sm		
2	rp		
8	hr		
1	nf	1	=nf
26	sk	7	=sk
1	sl		
1	nz		
20	rt	5	=rt
14	rs		
50	st	22	=st
1	sh		

30	tt	29	=tt
3	dd		
25	pp	8	=pp
28	kk	43	=kk
38	mm	1	=mm
41	nn	6	=nn
1	hh		
42	ll	9	=ll
12	ss	2	=ss
16	rr	3	=rr

BIBLIOGRAPHY

Allen, W. Sydney. 1973. *Accent and Rhythm: Prosodic Features of Latin and Greek: A Study in Theory and Reconstruction*. Cambridge University Press.

Arnott, D. W. 1970. *The Nominal and Verbal Systems of Fula*. Oxford University Press, New York.

Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. MIT Press, Cambridge.

Austerlitz, Robert. 1965. *Finnish Reader and Glossary*. (Second printing with corrections.) Indiana University Uralic and Altaic Series, v. 15. Bloomington.

Borowsky, Toni J. 1986. *Topics in the Lexical Phonology of English*. Graduate Student Linguistics Association Publications.

Bright, William. 1957. *The Karok Language*. University of California Publications in Linguistics, University of California Press, Berkeley.

Broman, C. and L. Goldstein. 1986. Towards an Articulatory Phonology. *Phonology Yearbook* 3, p. 219-252.

Broman, C. and L. Goldstein. 1987. Tiers in Articulatory Phonology, with Some implications for Casual Speech. *Haskins Lab, Status Report, Speech Research*. (Also in J. Kingston and M. Beckman, (eds) *Papers in Laboratory Phonology* 1, Cambridge University Press.

Broman, C. and L. Goldstein. 1989. Gestural Structures and Phonological Patterns. *Haskins Lab, Status Report, Speech Research 92.*

Broselow, Ellen. 1982. On Predicting the Interaction of Stress and Epenthesis. *Glossia* 16, p. 115.132.

Broselow, Ellen. 1992. Parametric Variation in Arabic Dialect Phonology. In Broselow, Ellen, Mushira Eia, and John McCarthy (eds.) *Perspectives on Arabic Linguistics: Papers from the Forth Annual Symposium on Arabic Linguistics* Johns Benjamins Publishing

Company, Philadelphia. p. 7-45.

Brown, R. and C. Hanlon. 1970. Derivational Complexity and the Order of Acquisition of Child Speech. In J. R. Hayes, ed., *Cognition and the Development of Language*. Riley, New York.

Chomsky, Noam and Morris Halle. 1968. *The Sound Pattern of English*. Harper and Row, New York.

Clements, George N. 1986. Compensatory Lengthening and Consonant Gemination in LuGanda. In Wetzels, Leo, and Engin Sezer (eds.) *Studies in Compensatory Lengthening*. Foris Publishers, Dordrecht.

Clements, George. N. and S. J. Keyser. 1983. *CV Phonology: A Generative Theory of the Syllable*. MIT Press, Cambridge.

Crowhurst, Megan. 1991. *Minimality and Foot Structure in Metrical Phonology and Prosodic Morphology*. Doctoral Dissertation, University of Arizona at Tucson.

Dixon, Richard M. W. 1979. Wargamay. In Dixon, Richard M. W. and Barry J. Blake (eds.) *Handbook of Australian Languages*, John Benjamins B. V., Amsterdam. Volume 2, p. 1-181.

Dixon, Richard M. W. 1988. *A Grammar of Boumaa Fijian*. University of Chicago Press, Chicago.

Dyen, I. 1965. *A Sketch of Trukese Grammar*. An essay of the American Oriental Society, New Haven.

Eades, Diana. 1979. Gumbaynggir. In Dixon, Richard M. W. and Barry J. Blake (eds.) *Handbook of Australian Languages*, John Benjamins B. V., Amsterdam. Volume 1, p. 245-384.

Fudge, E. C. 1969. Syllables. *Journal of Linguistics* 5, p. 253- 286.

Givon, Talmy. 1979. Prolegomena to Any Sane Creology. In Hancock, Ian F. (ed.) *Readings in Creole Studies*, E. Story-Scientia P.V.B.A. Scientific Publishers, Belgium.

Haviland, John. 1979. Guugu Yimidhirr. In Dixon, Richard M. W. and Barry J. Blake (eds.) *Handbook of Australian Languages*, John Benjamins B. V., Amsterdam. Volume 1, p. 27-181.

Hayes, Bruce. 1979. The Rhythmic Structure of Persian Verse. *Edebiyat* IV.2, p. 193-242.

Hayes, Bruce. 1982. Extrametricality and Stress. *Linguistic Inquiry* 13, p. 227-276.

Hayes, Bruce. 1985. *A Metrical Theory of Stress Rules*. Garland Publishing, Inc. New York.

Hayes, Bruce. 1989. Compensatory Lengthening in Moraic Phonology. *Linguistic Inquiry* 20.2. p. 253-306.

Hayes, Bruce. 1990. Diphthongization and Coindexing. *Phonology* 7.

Hayes, Bruce. manuscript. *Metrical Stress Theory: Principles and Case Studies*. Univerity of California, Los Angeles.

Hewitt, Mark. 1992. *Vertical Maximization and Metrical Theory*. Doctoral dissertation, Brandeis University, Waltham, MA.

Hyman, Larry M. 1970. The Role of Borrowing in the Justification of Phonological Grammars. *Studies in African Linguistics* 1, p. 1- 48.

Inkelas, Sharon and Young-Mee Yu Cho. 1993. Inalterability as Prespecification. *Language* 69.3, p.529-574.

Itô, Junko. 1986. *Syllable Theory in Prosodic Phonology*. Graduate Student Linguistics Association Publications, Amherst, MA.

Itô, Junko. 1989. A Prosodic Theory of Epenthesis. *Natural Language and Linguistic Theory* 17, p. 49-73.

Itô, Junko, and Armin Mester. 1992. *Weak Layering and Word Binarity*. Linguistic Research Center publication LRC-92-09, University of California, Santa Cruz.

Itô, Junko, Armin Mester, and Jaye Padgette. 1993. NC. Talk given at the Rutgers Optimality Workshop 1.

Karlsson, F. 1983. *Finnish Grammar*. Translated by Andrew Chesterman. Porvoo, Werner Soderstrom, Osakeyhtio.

Karttunen, Frances. 1970. *Problems in Finnish Phonology*. Indiana University Ph. D. thesis.

Kaye, Johnathan. 1990. 'Coda' Licensing. *Phonology* 7, p. 301-330. Press Syndicate of University of Cambridge, New York.

Kiparsky, Paul. 1982. Lexical Morphology and Phonology. In I-S Yang (ed.) *Linguistics in the Morning Calm*, Hanshin, Seoul.

Kiparsky, Paul. 1985. Some Consequences of Lexical Phonology. In Ewen, C. and J. Anderson (eds.) *Phonology Yearbook* 2, Cambridge University Press.

Kiparsky, Paul. 1993. Variable Rules. Talk given at the Rutgers Optimality Workshop 1.

Koefoed, Gert. 1979. Some Remarks on the Baby Talk Theory and the Relexification Theory. In Hancock, Ian F. (ed.) *Readings in Creole Studies*, E. Story-Scientia P.V.B.A. Scientific Publishers, Belgium.

Krauss, Michael. 1985. *Yupik Eskimo Prosodic Systems: Descriptive and Comparative Studies*. Alaskan Native Language Center, Fairbanks.

Kroger, Paul R. 1989. Discontinuous Reduplications in Vernacular Malay. *Proceedings of the Fifteenth Annual Meeting of the Berkeley Linguistics Society*. P. 314-329. Berkeley Linguistics Society, Berkeley.

Kueh, D. P. and K. L. Moll. 1976. A Cineradiographic Study of VC and CV Articulatory Velocities. *Journal of Phonetics* 4, p. 303-320.

Lamontagne, Gregory. 1993. *Syllabification and Consonant Cooccurence Conditions*. Graduate Student Linguistics Association Publications, University of Massachusetts, Amherst, MA.

Lamontagne, Gregory. and Tim D. Sherer. 1993. Optimality, Alignment, and English Level-Ordering Effects. Talk given at the Rutgers Optimality Workshop 1.

Lazard, Gilbert. 1992. *A Grammar of Contemporary Persian*. Translated by Shirley A. Lyon Mazda Publishers, New York.

Lightner, Theodore M. 1971. On Swadesh and Voegelin's 'A Problem in Phonological Alternation'. IJAL XXXVII. p. 27-237.

Lombardi, Linda. 1991. *Laryngeal Features and Laryngeal Neutralization*, Graduate Linguistics Student Association Publications, Amherst, MA.

Lowenstamm, Jean. 1987. Prosodic Government. Manuscript, University of Quebec, Montreal.

Lowenstamm, Jean and Jonathan Kaye. 1985. Compensatory Lengthening in Tiberian Hebrew. In Wetzels, Leo, and Engin Sezer (eds.) *Studies in Compensatory Lengthening*. Foris Publishers, Dordrecht.

McCarthy, John. 1979a. *Formal Problems in Semitic Phonology and Morphology*. Doctoral Dissertation, MIT, Cambridge, MA.

McCarthy, John. 1979b. On Stress and Syllabification. *Linguistic Inquiry* 10, p. 443-66.

McCarthy, John. 1985. OCP Effects: Gemination and Antigemination. *Linguistic Inquiry* 17, p. 207-263.

McCarthy, John. 1993. The Paralell Advantage. Talk given at the Rutgers Optimality Workshop 1.

McCarthy, John and Alan Prince. 1986. *Prosodic Morphology*. Manuscript, Amherst, and Waltham, MA.

McCarthy, John and Alan Prince. 1988. Quantitative Transfer in Reduplicative and Templatic Morphology. *Linguistics in the Morning Calm* 2, Linguistic Society of Korea (ed). Hanshin Publishing Co., Seoul. P. 3-35.

McCarthy, John and Alan Prince. 1990. Foot and Word in Prosodic Morphology: The Arabic Broken Plurals. *Natural Language and Linguistic Theory* 8, p. 209-282.

McCarthy, John and Alan Prince. 1993a. *Prosodic Morphology I: Constraint Interaction and Satisfaction*. Manuscript, University of Massachusetts, Amherst, and Rutgers University, New Brunswick, NJ.

McCarthy, John and Alan Prince. 1993b. Generalized Alignment. Manuscript, University of Massachusetts, Amherst, and Rutgers University, New Brunswick, NJ.

McCarthy, John, and Alison Taub. 1991. Review of Carole Paradis and Jean-Francois Prunet, The Special Status of Coronals. *Phonology,* p. 231-240.

McCawley, James. 1969. Length and Voicing in Tübatulabal. *Proceedings of the Chicago Linguistics Society* 5, p. 407-415.

Mester, Armin. 1986. *Studies in Tier Structure*. Graduate Student Linguistics Association Publications, Amherst, MA.

Myers, Scott. 1985. The Long and the Short of it: Vowel Quality Alternations in English. *Chicago Linguistics Society*.

Myers, Scott. 1987. Vowel Shortening in English. *Natural Language and Linguistic Theory* 5, 485-518.

Myers, Scott. 1991. Persistent Rules. *Linguistic Inquiry* 22, p. 315-344.

Padgett, Jaye. 1991. *Stricture in Feature Geometry*. Graduate Student Linguistics Association Publications, Amherst, MA.

Paradis, Carole. 1988. On the Theory of Constraints and Repair Strategies. *The Linguistic Review* 6, p. 71-91.

Paradis, Carole. 1992 (1986). *Lexical Phonology and Morphology: The Nominal Classes in Fula*. Garland, New York.

Paradis, Carole, and Jean-Francois Prunet. 1992. *The Special Status of Coronals: Internal and External Evidence*. Academic Press.

Pinker, Steven. 1979. Formal Models of Language Learning. *Cognition* 7, p. 217-283.

Pinker, Steven. 1989. *Learnability and Cognition*. MIT Press, Cambridge.

Prince, Alan. 1980. A Metrical Theory for Estonian Quantity. *Linguistic Inquiry* 11.3, p. 511-562.

Prince, Alan 1984. In *Language Sound Structure: Studies in Phonology Presented to Morris Halle by his Teacher and Students*, M. Aronoff and R. T. Oehrle (eds). MIT Press, Cambridge.

Prince, Alan. 1989. Metrical Forms. In *Phonetics and Phonology 1: Rhythm and Meter*, Paul Kiparsky and Gilbert Youmans, eds. Academic Press, Boston. p. 45-80.

Prince, Alan. 1990. Quantitative Consequences of Rhythmic Organization. *Proceedings of the Chicago Linguistics Society* 26-II, p. 471-490.

Prince, Alan and Paul Smolensky. 1991a. Optimality. Talk given at Arizona Phonology Conference. University of Arizona, Tucson.

Prince, Alan and Paul Smolensky. 1991b. Notes on Connectionism and Harmony Theory in Linguistics. Technical Report CU-CS-533-91. Department of Computer Science, University of Colorado, Boulder.

Prince, Alan and Paul Smolensky. 1992. Optimality: Constraint Interaction in Generative Grammar. Revised and augmented handout from talk delivered at WCCFL.

Prince, Alan and Paul Smolensky. 1993. Optimality Theory: Constraint Interaction in Generative Grammar. Manuscript. Rutgers University and University of Colorado at Boulder.

Rubach, J. and G. E. Booij. 1991. Edge Constituent Effects in Polish. *Natural Language and Linguistic Theory* 8, p. 427-463.

Saltarelli, Mario. 1970. *A Phonology of Italian in Generative Grammar*. Mouton, The Hague.

Sapir, David. 1965. *A Grammar of Diola Fogny*. Cambridge University Press, Cambridge.

Selkirk, Elisabeth. 1978. The Syllable. In H. van der Hulst and N. Smith (eds), *The Structure of Phonological Representations* Part II, Foris, Dordrecht.

Selkirk, Elisabeth. 1981. Epenthesis and Degenerate Syllables in Cairene Arabic. *Theoretical Issues in the Grammar of Semitic Languages*, Hagit Borer and Youssef Aoun (eds). MIT Department of Linguistics, Cambridge, MA. P. 209-232.

Selkirk, Elisabeth. 1990. A Two Root Theory of Length. *Papers in Phonology*, Elaine Dunlap and Jaye Padgett, eds. University of Massachusetts Occasional Papers in Linguistics 14. p. 123-171.

Selkirk, Elisabeth. 1992. Constraints on Syllable and Mora Structure. Handouts from lectures given at Stuttgart.

Selkirk, Elisabeth. 1993. The Prosodic Structure of Function Words. Talk given at the Rutgers Optimality Workshop 1.

Siegal, Dorothy. 1974. *Topics in English Morphology*. Doctoral dissertation, MIT, Cambridge.

Slobin, D. 1972. Children and Language: They Learn the Same Way All Around the World. *Psychology Today* 6, p. 71-82.

Smolensky, Paul. 1993. Harmony, Markedness, and Phonological Activity. Talk given at the Rutgers Optimality Workshop 1.

Sohn, Ho-min. 1976. *Woleaian Reference Grammar*. University of Hawaii Press, Honolulu.

Stampe, David. 1969. The Acquisition of Phonetic Representation. *Papers from the Fifth Regional Meeting of the Chicago Linguistics Society*. P. 41-56.

Stampe, David. 1973. *A Dissertation in Natural Phonology*. Doctoral Dissertation, NY: (1979, Garland Publishing Co.

Steriade, Donca. 1982. *Greek Prosodies and the Nature of Syllabification Processes*. MIT Doctoral Dissertation.

Steriade, Donca. 1991. Moras and Other Slots. *Proceedings of the Formal Linguistics Conference of the Midwest*, volume 1, p. 254- 280.

Swadesh, Morris and C. F. Voegelin. 1939. A Problem in Phonological Alternation. *Language*, volume 15.1. p. 1-10.

Teeter, Karl, V. 1964. *The Wiyot Language*. University of California Publications in Linguistics, University of California Press, Berkeley.

Thrainsson, Hoskuldur. 1978. On the Phonology of Preaspiration. *Nordic Journal of Linguistics* 1, p. 3-54. Universitetsforlaget, Boston.

Tranel, Bernard. 1991. CVC Light Syllables, Geminates and Moraic Theory. *Phonology* 8, p. 291-302.

Tyler, Stephen A. 1969. *Koya: An Outline Grammar (Gommu Dialect).* University of California Publications in Linguistics, University of California Press, Berkeley.

Voegelin, Charles F. 1935. *Tübatulabal Grammar*. University of California Publications in American Archaeology, Vol. 34, No. 2, p. 55-190.

Wexler, Kenneth, and Peter W. Culicover. 1980. *Formal Principles of Language Acquisition*. MIT Press, Cambridge.

Woodbury, Anthony. 1985. Graded Syllable Weight in Central Alaskan Yupik Eskimo (Hooper Bay-Cevek). *International Journal of American Language* 51, p. 620-623.

Woodbury, Anthony. 1987. Meaningful Phonological Processes: a Study of Central Alaskan Yupik Eskimo Prosody. *Language* 63, p. 685-740.

Yip, Moira. 1991. Coronals, Consonant Clusters and Coda Conditions. In Paradis, Carole and J.-F. Prunet, (eds.), *The Special Status of Coronals*, Academic Publishers.

Yip, Moira. 1993. Coda Constraints and Conditions on Syllable Weight. Manuscript, University of California, Irvine.

Zec, Draga. 1988. *Sonority Constraints on Prosodic Structure*. Stanford Ph.D. thesis.

Zec, Draga. 1993. Patterns of Gemination and Consonant Loss: Pali, Japanese, and Cross-linguistic. Talk given at the Rutgers Optimality Workshop 1.

Zvelebil, Kamil. 1970. *Comparative Dravidian Phonology*. Janua Linguarum Series Practica 80. Humanities Press, New York.

Sherer, Tim D. 1994 Prosodic Phonotactics. Available from GLSA Publications, Department of Linguistics, South College, University of Massachusetts, Amherst, MA 01003.

PREFACE TO THE COMPUTER-FILE VERSION

This is the first of six chapters of my dissertation. This version of my dissertation is divided into seven chapters, numbered 0 to 6. Chapters 1 through 6 contain the corresponding chapters of my dissertation. Chapter 0 contains the front matter, as well as the appendix and the sources cited.

This version of my dissertation was taken from the Graduate Linguistics Student Association (GLSA) version, which had several typos corrected over the deposit version. This version differs from the GLSA version with respect to some diagrams. Mainly in tableaux, I have substituted bracketing and subscripted morae for tree diagrams in representing syllables, as below:

$\{C\ V_{\mu}\}\ \{C\ V_{\mu}\}$	=	2 cv syllables
$\{C\ V_{\mu\mu}\}$	=	a syllable with a long vowel
$\{C\ V_{\mu}\ C_{\mu}\}$	=	a bi-moraic closed syllable
$\{C\ V_{\mu}\ C\}$	=	a mono-moraic closed syllable
$\{C\ V_{\mu}\ \{\ C_{\mu}\ \}\ V_{\mu}\ \}$	=	2 syllables with a geminate consonant

There are still some diagrams in their original forms. I hope that these are recoverable from context.

The pagination is different from the original version. Each chapter is paginated separately. There is a table of contents for each chapter at the beginning of that chapter.

This chapter serves as a general introduction to the dissertation. Since this one of the earlier works employing Prince and Smolensky’s 1993 Optimality manuscript, I include an introduction to the theory in this chapter.

Tim D. Sherer
TDSherer@Delphi.Com
December, 1994

INTRODUCTION

CHAPTER 1

INTRODUCTION

1.1 Licensing

Consider the following diagram, where "C" stands for consonant, "V" for vowel, and "σ" for syllable.

1) σ_1 σ_2

V_1 C_1 C_2 V_2

If we assume Prosodic Licensing (Itô 1986), which requires that every phonological element must be associated to an element of prosody, each of the segments, V_1, C_1, C_2, and V_2, must be associated to an element of prosody, such as a syllable (σ) or mora (μ). Elements that are not associated to prosody are subject to stray-erasure in Itô's system and very similar condition in the system adopted here. Following very standard assumptions about syllable structure, the vowels V_1 and V_2 in (1) are licensed as the nuclei of syllables, here σ_1 and σ_2. The consonant C_2 would be licensed as an onset of σ_2 in any language. Assume that C_1 and C_2 do not make up a complex onset. The purpose of this dissertation is to discover the principles behind various ways of licensing C_1 and the conditions under which C_1 cannot be licensed.

A consonant sequence is two or more consonants that are not separated by a vowel at the surface and that are not part of a complex onset. The bulk of this dissertation is devoted to consonant sequences over syllable boundaries.

Returning to a part of the diagram in (1), consider the case where there are two consonants that are adjacent to each other at the level of underlying representation.

2) V_1 C_1 C_2 V_2

This underlying representation will be mapped to a surface form. There are several possible results of such an underlying representation.[1] One possibility is that a language may epenthesize a vowel between C_1 and C_2, as in the following diagram, where "☐" indicates the epenthesis site

[1]This discussion follows Prince and Smolensky (1993), especially chapter 6 in predicting various syllable types.

of a vowel.

```
3)      σ1   σx      σ2
        |     /|       /|
        μ1  / μx     / μ2
        |   /  |     /   |
        V1 C1  □   C2  V2
```

The consonant C_1 is now licensed as an onset. This epenthetic vowel separates the two consonants at the surface. This approach is couched in terms that specify the location of an epenthetic vowel, leaving the nature of the vowel to another part of the grammar (see Selkirk 1981, Broselow 1982). There is now a vowel intervening between the two consonants, giving the surface string "VCVCV".

Another possibility is offered below, where angled brackets indicate that the consonant is not licensed in this configuration in this language.

```
4)          σ1                  σ2

            μ1                  μ2

            V1  <C1>       C2   V2
```

The consonant is not realized on the surface. In Itô's terms, it is stray-erased at or before phonetic implementation. Again, the sequence of two consonants does not occur on the surface. The surface string "VCV" is the result.

If we assume that the failure of C_1 to be licensed by σ_1 is general across the language and not peculiar to a subset of the consonants of the language, the two types of languages described above would be CV languages, lacking any consonant sequences. The first language described has an epenthetic vowel between the two consonants and the second fails to license the first consonant.

We may say, then, that having consonant sequences depends on licensing C_1 in some manner other than as an onset. One such licensing is shown below.

```
5)                σ1                  σ2

                  μ1  μx              μ2

                  V1  C1         C2   V2
```

Here C_1 is licensed by association to a mora. This is the method of licensing used in English, Arabic, Latin, and other languages. Because all morae are assumed to be the same,[2] we expect σ_1 to be heavy for purposes of stress, poetics, and templatic morphology. That is, σ_1 would count as equal to [CV:] for these properties.

Another licensing situation, which can be taken as a specific version of the previous one, is shown below, where the two consonants are now a single element of melody, or melodeme, abbreviated "G" in this diagram, associated to two syllables. This is the representation of a geminate or long consonant.

6)

σ_1			σ_2
μ_1	μ_x		μ_2
V_1		G	V_2

This situation differs from the one in (5). In (5), C_1 is dependant on μ_x for its licensing and would not be realized without μ_x. In (6), the melodeme G is licensed as the onset of σ_2. However, geminates and simple consonants do contrast in many languages and the cases in (5) and (6) do share the attribute of having a second mora in σ_1. While having a second mora in (5) permits C_1 to be licensed, having a second mora in (6) permits a geminate, instead of a simple consonant, to exist, even though the mora does not exclusively license the geminate's features.

Finally, consider the following representation, where there is no mora intervening between C_1 and the syllable node.

7)

σ_1			σ_2
μ_1			μ_2
V_1	C_1	C_2	V_2

This is what has been called a syllable appendix or simply an appendix. Appendixes have played a role in phonological theories (Fudge 1969, Selkirk 1978, Borowsky 1986, among others), although the exact formulation of the notion varies from theory to theory. Under the assumption implicit in the diagram in (7), we expect appendixes to be rather like onsets, because they are non-moraic consonants. We do not expect appendixes to participate in stress assignment or to interact with vowel length because these processes are stated in terms of morae (see hypercharacterized syllables, below, and in chapter 2). Appendixes are discussed at greater length in Chapter 5, as well as employed in Chapters 3 and 4.

The ways of licensing consonants that have been discussed are summarized below.

---

[2]But compare this assumption with Hayes (manuscript), especially chapter 7, for a view where not all morae are treated alike with respect to all elements of prosody.

8) Licensing Structures

MORAIC	CONSONANT	APPENDIX
MORAIC CODA	GEMINATE	
σ_1 σ_2	σ_1 σ_2	σ_1 σ_2
μ_1 μ_x μ_2	μ_1 μ_x μ_2	μ_1 μ_2
V_1 C_1 C_2 V_2	V_1 G V_2	V_1 C_1 C_2 V_2

I argue in this dissertation that the representations in (8), together with the possibility of adjoining to consonant "higher up" the prosodic hierarchy to a word node (see Chapter 5 for such proposals), are both necessary and sufficient for licensing C_1 and accounting for the distribution of geminates in natural language from a prosodic point of view. There are certainly other factors, principally the featural constraints required of adjacent consonants (Prince 1984, Yip 1989, 1991, Lamontagne 1993; this work rests on Obligatory Contour Principle effects, see McCarthy 1985, Mester 1986) as well as phonetic effects, like hiding phenomena (Broman and Goldstein 1986, 1987, 1989), but once we understand the principles governing the structures above, we have an explanation for why languages permit consonant sequences and geminates from a prosodic point of view.

An analysis of the consonant sequence and geminate facts from a prosodic point of view must describe not only the existence of these structures in one language, but must also describe the variation found from one language to the next. In other words, we need to be able to prohibit a structure in some languages and permit it in others. Furthermore, some of the structures above can be found in the same languages, so our analysis must permit more than one type of structure per language. For instance, there are geminates in some languages with consonant sequences, while other languages have geminates but no consonant sequences.

My analysis of these facts will be stated in Optimality Theory (Prince and Smolensky 1993), a system of interacting constraints. The different constraints will each be simple and independently motivated. In some permutations, they will give rise to certain structures, in others, to other structures. The goal here is to achieve great generality from simple elements in different relations to each other. Optimality Theory permits this generality because it is a theory of the interaction of phonological constraints. The constraints themselves are simple. The theory of interaction involves differences in the ranking of the constraints with respect to each other in individual grammars, where the notion of *ranking* is a strictly defined element of the theory of constraint interaction.

1.2 Hypercharacterized Syllables

The distinction between light and heavy syllables is widely attested. In some languages, it is possible for a syllable to be heavy in more than one way--by being closed or by having a long vowel. It is also possible for a syllable to more than qualify as heavy by having a long vowel and being closed. That is, while either a long vowel or a closing consonant would be sufficient, such syllables have both attributes.

These "doubly heavy" syllables are referred to by Allen (1972) as "hypercharacterized" syllables. His definition, couched in a motor theory of speech, is, "There is also the possibility of 'double arrest', i.e. both thoracic and oral in sequence (~V+C) [=...VVC; TS]" (Allen, p. 66.). I am generalizing the term *hypercharacterized syllable* slightly to include any syllable that has more that two elements in its "rime." The following are examples of hypercharacterized syllables.

9) Tri-moraic Hypercharacterized Syllables

a)		b)		c)	
	σ		σ		σ
	μ μ μ		μ μ μ		μ μ μ
	V C C		V C		V G

10) Mono- or Bi-moraic Hypercharacterized Syllables

a)		b)	
	σ		σ
	μ μ		μ
	V C		V C C

These syllables differ in their weight (2 versus 3 morae) and in the number of consonants in the syllable (1 versus 2). All of them are doubly "closed" in some sense.

1.2.1 Tri-moraic Syllables

Examples (9a)-(9c) are tri-moraic. Tri-moraic syllables are not permitted in a variety of languages. Beyond the cases where tri-moraic syllables are prohibited, there are also languages where the prohibition gives rise to alternations, as we can see in closed syllable shortenings and degemination after long vowels. In this approach, there is a prohibition against tri-moraic syllables, but that prohibition must interact with other constraints of the language, being unviolated in some languages, being violated in other languages.

Example (9a) might arise if a language prefers to keep its consonants and has no other way to incorporate C_2 into a syllable. Example (9b) will arise if both vowel length and the consonant must be maintained. Example (9c) will arise if vowel length and geminate length are both maintained. These preferences are stated precisely below as interacting, conflicting constraints, within Optimality Theory, which is taken up in 1.3.

Languages usually only make a two-way distinction between light (mono-moraic) syllables on the one hand and heavy (bi- and tri-moraic) syllables on the other, but some languages do make a three way distinction. Hindi places stress first on tri-moraic, then bi-moraic syllables (Hayes (manuscript) and sources cited there, compare with Ohala 1977). Persian counts tri-moraic syllables as a whole foot in its system of versification (Hayes 1979).

1.2.2 Non-tri-moraic Syllables

Examples (10a) and (10b) exploit another way of licensing consonants, the appendix. Essentially, these are languages where consonants are not moraic and a single consonant is permitted as appendixal (10a) or two are permitted (10b).

There is no special prohibition against syllables with a long vowel and an appendixal consonant. They will violate the prohibition against syllables closed with an appendix, but this is not on a par with the prohibition against tri-moraic syllables, which refers only to syllables with three morae, not to moraic consonants in general. In principle, a syllable could rack up a string of appendix consonants of any length. This doesn't happen, partly because of the dissimilatory effects that adjacent segments are subject to (Prince 1984, Yip 1989, 1991, Lamontagne 1993) and the possibility of simplifying sequences over time, partly through phonetic effects of hiding (Broman and Goldstein 1986, 1987, 1989).

1.2.3 Role of Hypercharacterized Syllables

Hypercharacterized syllables enjoy a special place in our pantheon of licensing. In hypercharacterized syllables, we are trying to license more material than will "fit" into a syllable. This is a conflict between two elements of phonology. Tri-moraic syllables are dispreferred, but will occur, for instance, if the language decides that it must keep the vowel length and maintain all consonants to the surface without employing epenthesis. Syllables with appendixes are dispreferred, under the assumption that "codas" are disfavored. Again, a language may choose to have such representations to maintain underlying representation to the surface. These preferences are formalized below as interactions of phonological constraints, not vague statements of the preferences of a grammar.

1.3 Optimality Theory

Optimality Theory is a system of constraint interaction in phonology. This system is laid out in Prince and Smolensky (1993). Optimality Theory was also employed in earlier work by Prince and Smolensky (1991a, 1991b, 1992), and employed in a variety of works, including McCarthy and Prince (1993a) and sources cited in Prince and Smolensky (1993), as well as a variety of subsequent work. For principles of Optimality Theory, I will cite Prince and

Smolensky (1993), even though some of these principles were presented or foreshadowed in earlier Prince and Smolensky work.

In this section, I will give an overview of Optimality Theory. First, I will give an abstract, overall generalization of the theory, then I will give a specific example. Next, I will turn to some notation. Finally, I will give some particular assumptions about underlying representation, some of which are not part of Prince and Smolensky's work, but which are consistent with Optimality Theory.

1.3.1 The Character of the Theory

The standard theory of phonology, to be found in Chomsky and Halle (1968) and many works that build on it, starts with an underlying form and maps it to a surface form by a series of rules, where the output of each rule serves as the input to the next. Each rule has a specific input and structural change. The result of the derivation is the surface form.

In analyzing language this way, certain phonological patterns emerge in the output of the phonological rules in this sequential system. For instance, many languages break up sequences of consonants under certain conditions. The rules that may be used to avoid undesired sequences (epenthesis, deletion, dissimilation) are varied. Yet all of these varied rules have the same effect. They prevent sequences of consonants from arising. We cannot say that the language is trying to avoid clusters. The languages in question simply have one or more of a variety of rules that have the effect of breaking up clusters.

Optimality Theory, proposed by Prince and Smolensky (1993), focuses not on the step by step derivation, but on the constraints that underlie phonological changes. Optimality Theory gives status to the various constraints on well-formedness that drive phonology. Once these constraints are recognized and allowed to interact directly, a major simplification of the derivational component is possible.

The course of an Optimality Theory derivation is as follows (see Prince and Smolensky 1993, especially chapters 1 and 2.) An underlying form is a morpheme or concatenation of morphemes, much as it is in any other theory of phonology. I will be making particular assumptions about the form of underlying representation in this dissertation, but these are not necessary parts of Optimality Theory. The underlying form is fed into a function called GEN which creates the set of candidates that undergo evaluation. GEN provides structure. It will associate segments to syllables and syllables to feet, as well as providing some associations lines of the feature geometry tree. In doing these things, GEN creates all possible representations for an underlying form. That is, GEN produces a great number of phonological forms. These forms are called *candidates* and they are the "possible" surface forms for the particular underlying representation in question. They are only possible forms in the sense that they, and no others, could be the output, even though many of them would never surface in any language.

The candidates are next turned over to the constraints of the grammar. Recognizing the centrality of output constraints in generative phonology, Optimality Theory uses a system of constraints to evaluate possible forms. The constraints are assumed to be universal either completely or as constraint schema to be set for particular languages.

These constraints are ranked. That is, the candidates are submitted to one constraint, then to the next. That is, all the candidates created by GEN are evaluated by one constraint, say X, then by another, say Y. Some meet X, while others violate it. Those that violate X are dropped from further evaluation. Those that do meet constraint X are passed to the next constraint, Y, where the procedure is repeated. Because X evaluates candidates before Y does, X may rule out candidates that would meet constraint Y. In the situation described here, X is ranked above Y.

When the constraints have ruled out all but a single candidate, that is the actual form to which the underlying form is mapped. This form is referred to as *most harmonic* or *optimal*. The form chosen may violate one or more of the constraints of the language. This happens in two situations. First, when all the candidates violate a constraint, they are passed to the next constraint for further evaluation, so that an optimal form may still violate some constraints. Second, once all the candidates except for the optimal one have been ruled out by the constraints, that optimal candidate may have additional violations of lower ranked constraints. The term *optimal* does not imply meeting every constraint.

This is the course of an Optimality Theoretic derivation. This description is very much a thumbnail sketch. Issues of this type are discussed at length in Prince and Smolensky (1993; see especially chapters 5 and 10). In the next sub-section, I will discuss these elements with an underlying form.

1.3.2 An Example

Let me provide an example along the lines of ones that will be discussed below. Suppose we have an underlying representation, as below, with long vowels indicated by two superscripted morae and short vowels with a single mora.[3]

11) /$ta^{\mu\mu}ktu^{\mu}$/

This could be a simplex form or the result of concatenation. GEN makes no distinction between different sources of inputs, although specific constraints may treat monomorphemic forms differently than polymorphs (see McCarthy and Prince 1993a, 1993b on Alignment constraints). Whatever its source, this form would be fed into GEN. All potential surface forms would be created, including many that would never occur in any language. Some candidates are shown below.

[3]This notation will be further explained below. For the time being, simply consider the morae to be pre-linked to the melodemes in question.

12) Sample Candidates for $/ta^{\mu\mu}ktu^{\mu}/$

a) σ σ μ μ μ μ t a k t u	b) σ σ μ μ μ t a k <t> u	c) σ σ μ μ μ t a <k> t u
d) σ σ μ <μ> μ μ t a k t u	e) σ σ σ μ μ μ μ t a k ☐ t u	f) σ σ μ μ μ μ t a k <t> u
g) σ <σ> μ μ <μ> t a <k t u>	h) σ σ μ μ μ t a k t u	

Each of these candidates has a different structure, although candidates (12a) and (12h) have identical phonetic forms. This example is not from any language. In any language with an underlying form $/ta^{\mu\mu}ktu^{\mu}/$, these candidates will be among the forms evaluated. There are, however, a great deal of other candidates available for evaluation. I chose these candidates as examples because they vary with respect to how the sequence of two consonants is incorporated into the structure or not incorporated, along the lines of the discussion earlier in this chapter. I will describe the candidates one at a time, both to explain different sorts of candidates and to explain any notation that may be unfamiliar.

13a) /k/ associated to syllable via mora

b) /t/ not associated to prosody; "unlicensed" or "stray erased"

c) /k/ not associated to prosody; "unlicensed" or "stray erased"

d) /k/ associated to syllable via a mora; a vowel's mora is left unparsed and the vowel is short in this candidate

e) /k/ is parsed as the onset of an empty syllable, where ☐ signifies that there is no phonetic content

f) /k/ is a geminate, /t/ is unlicensed

g) /k/ and entire second syllable left unparsed

h) /k/ parsed as an appendix consonant

Some of these candidates are the actual surface forms for underlying representations like /ta$^{\mu\mu}$ktu$^{\mu}$/ in different languages. Candidate (13a) is a hypercharacterized syllable, which is permitted in some languages, but not others. Candidates (13b) and (13c) are what would be called deletions of consonants in process oriented systems. Candidate (13d) is a "closed syllable shortening", familiar from Arabic, English and Yawlemani. Candidate (13e) is an "epenthesis" of a type familiar from dialects of Arabic and many other languages. Candidate (13g) is a truncation, similar to truncations found in Lardil. Candidate (13h) is a closed syllable in a language without moraic consonants.

Candidate (13f) has the first consonant licensed as moraic, while the second consonant is unlicensed. I know of no language which employs (13f), nor would we expect to find a language which employed (13f). We will return to this point below.

The constraints which evaluate the candidates are assorted well-formedness conditions on phonological representation, as noted above. These constraints are universal, either in their full form or universal in the sense that there is a schema for a particular constraint that has particular values set for one grammar and may have other values for another grammar (Alignment constraints are of this form, see McCarthy and Prince 1993a, 1993b).

The constraints interact by being ranked. A grammar will have a set of constraints ranked from highest to lowest. The candidates are evaluated by the highest constraint, then the next highest constraint then the next highest constraint, down the line. When a constraint rules out a candidate, it is permanently out of the running, with out any notion of repair, as in some other theories involving constraints (Myers 1987a, 1987b, 1991, Paradis 1988).

Candidates may tie with respect to a constraint, either meeting it or violating it. In the event that two or more candidates violate a constraint and there are no other candidates, the candidates move on to the next constraint, until one candidate remains.

For an example of ranking and evaluation, we return to the forms in (12). These candidates and many others will be fed into the constraints. Imagine that there are five constraints. These constraints are due ultimately to Prince and Smolensky (1993). These constraints will be introduced in more detail in chapter 2.

14a)	$*\sigma_{\mu\mu\mu}$	Syllables are maximally bi-moraic.
b)	FILL_{μ}	Mora must have phonological content.
c)	*APPENDIX	Appendix consonants are not allowed.
d)	$\text{PARSE}_{\text{SEGMENT}}$	Segments must be parsed.
e)	PARSE_{μ}	Morae must be parsed.

These are not new principles. That syllables must not have more than two morae has been suggested as a universal by Lowenstamm and Kaye (1985), McCarthy and Prince (1986), Lowenstamm (1987), Steriade (1991), Paradis (1992 (a translation of a 1986 work)), but has roots in earlier work in closed syllable shortening, for instance Myers (1985, 1987); also of note is Borowsky (1986), who showed that the two mora limit is relevant in English, a language where it appears to be utterly false.

The constraint against empty morae follows from the approach to epenthesis taken by Selkirk (1981) and Broselow (1982) where the location of an epenthetic vowel is specified from principles of syllabic structure. This constraint simply explains why epenthesis does not happen all the time; because there is a violation of this constraint, which is a cost to the grammar.

The constraint *APPENDIX is one of two constraints used in this dissertation to prefer open syllables. It is a form of Prince and Smolensky's constraint NOCODA (Prince and Smolensky 1993, chapter 2). That certain languages will not permit codas, however, is a long standing fact, included as a descriptive generalization about many languages. The innovation of Prince and Smolensky is to integrate the fact about codas into a system that accounts for the full range of phonological phenomena, such as epenthesis and deletion.

The two PARSE constraints require that different elements of the representation must be associated to higher levels of prosodic structure. This is a form of Itô's (1986) Prosodic Licensing requirement. That is, any time that something cannot be integrated into the structure of the word, this principle is violated. Such elements are not realized on the surface--in Itô's terms, they are stray erased.

These constraints, then, are not new to Optimality, although Optimality does permit us to state directly some of the generalizations that have long driven the rules of phonology. That is, *APPENDIX is now a principle of grammar, not an implicit driving force behind rules that prevent segments from being codas.

What is new about Optimality Theory is that it is a theory of the interaction of constraints. To see how these constraints interact, we will return to the candidates in (12). In the following diagram, called a *tableau*, the candidates are evaluated by the constraints.

15) Example Tableau

	*APPENDIX	$*\sigma_{\mu\mu\mu}$	$PARSE_{SEG}$	$PARSE_{\mu}$	$FILL_{\mu}$
a		*!			
b			*!		
c			*!		
d				*!	
☞ e					*
f		*!	*		
g			*!**	*	
h	*!				

The first constraint is violated only by candidate (15h). This violation will forever remove this candidate from the evaluation procedure. The "!" points out such a fatal violation on a tableau. Cells that follow a fatal violation are shaded, since they will not be relevant to any following evaluation. Note also on this point that (15h) does not violate any other constraints here, while the optimal candidate might violate one or more lower constraints. This does not matter--when a candidate is removed from consideration, that removal is permanent.

Moving across the tableau to the right, the next constraint is $*\sigma_{\mu\mu\mu}$, which removes (15a) and (15f) from consideration. Next, (15b), (15c) and (15g) are out by $PARSE_{SEGMENT}$. Candidate (15g) incurs multiple violations of $PARSE_{SEGMENT}$.

$PARSE_{\mu}$ removes (15d). Candidate (15g) also violates $PARSE_{\mu}$, but that violation is not important since (15g) is already out.

The only remaining candidate is (15e), which is in violation of $FILL_{\mu}$ and nothing else.

The constraints, taken as a whole, will have two effects. In some ranking relationships, they will map the underlying form to some output and in other ranking relationships, to other outputs. Grammars differ in having different ranking relationships and consequently different outputs. Below are the same constraints and the same candidates in another tableau. The only difference between this and the earlier tableau is that the constraints are in a different ranking relationship.

16) Example Tableau

	*APPENDIX	$FILL_{\mu}$	$PARSE_{\mu}$	$PARSE_{SEG}$	$*\sigma_{\mu\mu\mu}$

☞	a					*
	b				*!	
	c				*!	
	d			*!		
	e		*!			
	f				*!	*
	g			*!	***	
	h	*!				

The output of this second tableau is candidate (15a), not candidate (15e), as in the earlier tableau. These different possibilities of ranking of constraints are the main and perhaps the only way in which grammars differ in Optimality Theory. This attribute of Optimality Theory is taken up at length in Chapter 2, below, building on work of Prince and Smolensky (1993, chapter 6).

The other effect of the constraints as a group is to rule out certain forms from any language, regardless of the ranking of the constraints. Prince and Smolensky (1993) show this effect with the syllable parse "VC.V", where a single intervocalic consonant is parsed as a coda. This parsing never occurs, even in languages where codas are permitted and Prince and Smolensky (1993) derive this fact from several constraints in any ranking relationship whatever.

In the example here, candidates (15f) and (15g) will never be the optimal ones from this set of constraints. Candidate (15f) incurs a violation of both $*\sigma_{\mu\mu\mu}$ and $\text{PARSE}_{\text{SEGMENT}}$, but there are always alternative candidates, here (15a), (15b), and (15c), which incur one less violation than (15f) of a relevant constraint. Candidates (15f) and (15a) tie on $*\sigma_{\mu\mu\mu}$, but (15a) does not violate $\text{PARSE}_{\text{SEGMENT}}$. We must bear in mind that there may, in principle, be other constraints that could force (15f) or (15g) to be optimal. Indeed, a candidate like (15g) is optimal in Lardil (Prince and Smolensky 1993, chapter 7) as the result of interactions with a constraint on word size.

This section has summed up the course of an Optimality Theory derivation. See Prince and Smolensky (1993), especially chapter 2, for discussion of the Optimality Theoretic mapping process. The ability to have different results follow from different constraint ranking relationships is an important point, taken up in the next section and at length in chapter 2, below.

1.3.3 Factorial Typology

Because different languages employ different rankings of the same set of constraints, any analysis of a phenomenon from one language is an implicit typology of possible languages. Suppose a language employs three constraints, A, B, and C, in the ranking relationship A >> B >> C (">>" indicates a ranking relationship, with A above B). We would also expect that other languages might employ one of the rankings below.

17) Other Ranking Relationships for A, B, and C

A	>>	C	>>	B
B	>>	A	>>	C
B	>>	C	>>	A
C	>>	A	>>	B
C	>>	B	>>	A

As shown above, we can expect different outputs from grammars with different ranking relationships. This is called a *factorial typology* by Prince and Smolensky (1993; see especially chapters 3 and 6). It does not follow that there will be six possible outputs from this interaction of three constraints, one from each ranking. Several of the rankings might be non-distinct; those rankings might produce the same output. All rankings are predicted to occur, however, within some limits. Chapter 2 is devoted to a factorial typology using some of the constraints discussed above.

1.3.4 A Note on Tableaux

A tableau is a representation of some of the candidates created by GEN for an underlying form and some of the constraints evaluating them. It is intended to be a graphic representation, where the reader can easily follow the interaction of constraints (see Prince and Smolensky 1993; tableaux are introduced in chapter 2). A tableau shows which surface candidate a particular underlying form will map to. In this sense, it is rather like a truth table from formal logic. That is, a tableau is evaluated according to strict principles of Optimality Theory to provide an output for any input,
just as a truth table is evaluated according to the principles of formal logic to provide a truth value as an output for any input.

I have already introduced the notation employed in tableaux. I will simply sum it up here. Below, I give an example, derived from one given above.

18) Example Tableau

		*APPENDIX	$*\sigma_{\mu\mu\mu}$	$\text{PARSE}_{\text{SEG}}$	PARSE_{μ}	FILL_{μ}
	a		*!			
	b			*!		
	c			*!		
	d				*!	
☞	e					*
	f		*!	*		
	g			*!**	*	
	h	*!				

Tableaux are organized with candidates in a column down the left side and constraints across the top, from left to right in descending rank. A violation of a constraint is marked with an asterisk. The violation which removes a candidate from the running is marked with an exclamation point. This is only to draw our attention to it and has no status. Once a candidate is out of the running, whether or not it violates constraints is irrelevant, a fact reflected on the tableau by having those lower cells shaded. A small left hand points to the optimal candidate. Again, this is only to draw our attention to it and has no status. Not every ranking can always be completely established. In this tableau, the constraints $*\sigma_{\mu\mu\mu}$ and $\text{PARSE}_{\text{SEGMENT}}$ are separated by a dashed line, indicating that their ranking with respect to each other cannot be established in the language in question. Technically speaking, either of the two violations that (18f) incurs could be the fatal one when the ranking of two constraint cannot be ascertained. However, I will continue to mark the left-more of two violations with "!" in cases like this.

The tableau above is what I call a *full tableau*, because it represents all of the candidates under discussion with respect to all of the constraints under discussion. In the following chapters, I use full tableaux at the end of the discussion of a language to show the entire analysis for that language.

It is useful, when analyzing a language, to break the discussion down to two constraints evaluating two candidates. Because these mini-tableaux help us discover the rankings of the constraints, I refer to these as *ranking tableaux*. A ranking tableau serves a different purpose in the exposition, but it is nothing more than a close-up of the full tableau.

Below is a ranking tableau taken from the full tableau above. When using a ranking tableau, we know what the output form for the underlying form of the language is. In this case, we know that it is (19e), as marked on the tableau.

19) Example Ranking Tableau

		$*\sigma_{\mu\mu\mu}$	$FILL_{\mu}$
	a	*!	
☞	e		*

Since we know that (19e) is the correct output form, we know that this must be the correct ranking relationship for these two constraints. If the ranking relationship were reversed, then (19e) would be ruled out, and some other candidate (possibly, but not necessarily (19a)) would be the actual output form. We know this because we know that at least (19a) would be more harmonic than (19e).

There is a caveat to this sort of argument. This argument assumes that there is no other constraint ranked above these two that would remove (19a) from consideration. We must find any possible constraint that could have this effect and establish its ranking if we are to have a valid ranking argument. In practice, this is not too difficult since the candidates that we evaluate usually differ from each other only in relevant ways. That is, (19a) and (19e) differ from each other only in the way in which /k/ is parsed and the existence of an empty mora. Such similarity limits the number of other constraints that we have to consider as possibly interfering with our determination of ranking.

1.3.5 Why Optimality?

What Optimality Theory contributes to our understanding of licensing is that syllable final consonants are disfavored, but sometimes must be kept, in order to satisfy other constraints of the language. When they are kept, it will be in the way that is least costly to the individual grammar as measured in the specific terms of the Optimality Theoretic formulation of the grammar. That is, when a consonant is parsed in violation of constraints against syllable final consonants, that consonant will be parsed in such a way as to violate only the lowest relevant constraint.

What Optimality Theory can contribute to our understanding of hypercharacterized syllables is, first and foremost, the notion that all languages have a constraint against tri-moraic syllables. Languages do not vary by either having or not having this constraint, in the nature of a parameter. This constraint, like all other constraints, has the potential to conflict with other constraints. In some languages, the constraint against tri-moraic syllables takes precedence over other constraints. In other languages, the other constraints take precedence over it. Or, in some languages, the constraint against tri-moraic syllables is violated by the optimal candidate of some underlying representations, while that same constraint is not violated by the optimal candidate of other underlying representations, as in Finnish (see 6.2, below.)

Finally, Optimality Theory offers us a way to realize the goal of achieving complex results from the interactions of simple principles. We have the possibility of constraints in any possible ranking. Such a system creates different outputs for the same underlying representations in different languages while using the same set of simple constraints. A surprising amount of variation can be derived this way, all without the notion of a set of parameters with individual settings that vary form language to language.

1.3.6 The Nature of GEN and Underlying Representation

Two elements have been a part of all work in Optimality Theory to date. These are the principles of Freedom of Analysis and Containment. Both of these principles are descriptions of part of the way that GEN works. These are given below, quoted from McCarthy an Prince (1993a, p. 20), following Prince and Smolensky (1993, see chapters 1 and 2).

20a) Freedom of Analysis

Any amount of structure may be posited.

b) Containment

No element may be literally removed from the input form.
The input is thus contained in every candidate form.

Freedom of Analysis assures us that the candidate set is not constrained in any linguistic sense. No limits are imposed here. This seems, on first sight, to be useless, only providing excess candidates, many of which could not be optimal in any language. In fact, this approach to GEN permits us to remove a redundancy from linguistic theory. That is, on this view, well-formedness is evaluated only by the constraints. The same well-formedness conditions are not hidden in the generative process (rules and conditions on rule application), as they are in the standard approach.

In this dissertation, GEN can be characterized as producing all the candidates that fit the following schema.

21) Characterization of GEN

a)	Prosodic Word	→	Foot*	(C^*)
b)	Foot	→	σ^*	(C^*)
c)	σ	→	(C^*) μ^*	(C^*)
d)	μ	→	{ V, C }	

In addition to these rules, candidates will vary in whether the elements have association lines between them. That is, the following are both parts of possible outputs of GEN.

22) μ μ

C C

The rewrite arrow in this characterization of GEN does not imply linking, although it does imply co-indexation, as described below.

As for the particular rules offered here, (21a-c) contain X^* (that is, zero to infinite X's). Candidates with any number of, say, feet, which lack underlying content to fill them will constitute FILL violations (see Prince and Smolensky 1993,
for instance, chapters 6 and 8) and consequently be less favored than those that lack the empty structure. Putting an arbitrary limit on the number of feet in (21a) is unnecessary and redundant.

We might ask whether the lower limit on the elements in the right side of the rule should be one, not zero. That is, every element on the right side of the arrow is optional. Again, this is an attempt to keep GEN as simple as possible, on the assumption that specific constraints will favor headed constituents over headless ones (see Itô and Mester (1992) and Selkirk (1993)). On the specific point of (21c), there is at least one very credible analysis that calls for non-moraic syllables. This is a lengthening in Woleaian (Sohn 1975), which is triggered by the devoicing of a word final vowel. If the word final vowel is treated as non-moraic, the adjustment of the length of a preceding vowel is understandable.

Each of the rules (21a-c) has an optional consonant on the right side of the rule. This is the appendix consonant, mentioned above and discussed in 2.4 and 2.5.2, in the analyses in chapter 4, and discussed in chapter 5. Word-final appendixes have been attested in several languages (see, for instance, Selkirk (1978) and sources discussed in chapter 5, below), hence rule (21a). Non-moraic syllable final consonants are common enough, hence rule (21c). The status of the consonant in rule (21b) is not attested, but see chapter 5 for discussion of this point.

What is still missing from the rules in (21) and is tangential to this dissertation, is the question of associating, say, σ directly to Prosodic Word or μ to Foot. Such associations are considered disfavored by Prince and Smolensky (1993, chapter 3).

We might want to consider reducing (21) to a single schematic rule, much the same way that X-bar syntax reduces the various rules of phrase structure to a schema of bar levels. Something like the following rule might serve.

23) $X \rightarrow Y^* \quad (Z^*)$
Where X, Y, Z are members of the prosodic hierarchy, with
$X > Y$ and $Y > Z$.

Such a proposal would need to be worked out and I will not do so here.

Freedom of Analysis does not require that GEN create candidates with morae dominating syllables or with [+voice] over the root node. GEN is a function stated in an algebra of linguistic units. The rules in (21) are a crude attempt at understanding these principles. Such bizarre candidates as described in this paragraph will not result from that function, any more than a function which can only add natural numbers can produce a fraction or a negative number.

As for Containment, this attribute is vital if we are to capture generalizations about phonological alternations through the constraints. That is, if GEN is allowed to remove elements from the representation, producing candidates without certain violations, those candidates would be preferred. The constraints would not be able to attach a price to deletion because the constraints would never see signs of the deletions.

Another premise of this work dealing with underlying representation and not a necessary part of Optimality Theory itself is the representation of geminates. I am adopting the moraic theory of geminates (McCarthy and Prince 1988, Hayes 1989). This approach states that a geminate is represented in the lexicon as simple consonant with a mora. Similarly, vowel length is represented in the lexicon with morae. I follow Hayes (1989) and McCarthy and Prince (1993a), in having long vowels marked with two morae underlyingly and short vowels with a single mora underlyingly.[4]

We cannot say, however, that morae are pre-linked, if we assume Containment and that GEN is incapable of cutting association lines.[5] Such a move would eliminate the possibility of compensatory lengthening as a result of degemination. Consider the following underlying representation, with prelinked gemination.

24)		μ	
	V	G	V

The mora could be left unparsed, leaving the candidate without a surface geminate. That mora cannot be parsed to the vowel melodeme without also licensing the consonant.

25)	σ		σ
	μ μ		μ
	V	G	V

[4]McCarthy and Prince (1988) represent geminates with a single mora, long vowels with a single attached mora, and short vowels with no attached mora. A rule assigns a mora to every vowel, giving a short vowel its single mora and a long vowel its second mora. This approach removes some redundancy from the lexicon. Unfortunately, this move does not seem to be possible in Optimality Theory. If we were to adopt that, then GEN would have to add a mora to every vowel, which would involve complicating GEN.

[5]See Itô, Mester, and Padgett (1993) for a system that makes just the opposite assumption.

In order to avoid this problem, let us adopt the idea of Co-indexation from Zec (1993). Below I give Zec's generalized statement about Weight-by-Position in a Co-indexing and Projection framework (=Zec (3)).

26) Segment r_i projects a mora iff it is not followed by a more sonorous segment r_j.

μ_i

r_i r_j

We would propose that geminates are simple segments which are co-indexed with a mora underlyingly. Linking is a separate element of the theory. A segment is optimally linked to a mora which it is co-indexed with, but it is also possible for linking to some other element, such as a vowel in the case of compensatory lengthening.

We may extend this way of thinking to feature geometry, stating the underlying representation of forms in co-indexing rather than linking, following McCarthy (1993).[6] If we assume that feature geometry is provided completely or partially by GEN, we simplify the underlying form to contain only co-indexed features and morae.

The co-indexing relationship need not include domination or other tree-theoretic concepts. It can be, essentially, a set relationship, where all the features and morae that are co-indexed, belong to the same set.

Since there is no relationship of linking in underlying representation, we will have to constraint the way in which features and morae are linked in candidates. To see why this is true, look at the following diagram, which is a schematic representation of a short vowel followed by a geminate at underlying representation. This form was included in (8), but here the geminate and its mora are marked as co-indexed and here the linking relation is not marked.

27) Form with Geminate

σ_1 σ_2

μ_1 μ_i μ_2

V_1 G_i V_2

We can see that it is permissible to link the geminate to the mora, but we see that it is also

[6]Hayes (1990) uses a co-indexing theory to separate out the classificatory function of features (features grouping together) from other elements of feature geometry. While this is rather different from the co-indexing approach adopted here, it employs two separate relations, instead of one relation, linking.

possible to link that mora to the proceeding vowel. We do not want to absolutely prohibit such a linking, since it is necessary in some cases (see, for instance, Clements (1986) for a case where a vowel's mora would be analyzed as linking to another segment in exactly this way.) Instead, linking to some element to which another element is not associated must incur a violation. In this case, I am proposing the following constraint.

28) Segmental Integrity

Elements must be linked to elements to which they are co-indexed.

Perhaps there will have to be separate constraints for morae and for features or separate constraints for different features. The full ramifications of this constraint are beyond the scope of this dissertation. It does, however, arise in the course of the discussion below, from time to time.

1.4 Structure of this Dissertation

Chapter 2 will be a factorial typology involving some of the constraints introduced earlier in this chapter. There we will see the typological predictions for the structures in (8) and different types of hypercharacterized syllables. This system predicts a cross-classification of geminates and consonant sequences. A language may have one or the other or both or neither in this system because there is no connection between geminates and consonant sequences. Hypercharacterized syllables are restricted from occurring by the constraint against tri-moraic syllables, and also cannot occur in a language that lacks the prerequisite geminates or consonant sequences in the first place.

Another question dealt with in chapter 2 is the learnability of underlying representations and the learnability of constraint rankings based on the surface forms available to a language learner. There it will be suggested that certain ranking relationships are not stable, because the surface forms of the language lead language learners to re-rank the constraints.

Chapter 2 also deals with the question of how geminates may be derived by prosodic interactions and by the concatenation of morphemes.

Chapter 3 is an analysis of the hypercharacterized syllables of Finnish. Finnish has two varieties of medial hypercharacterized syllables in its morphemes. The first type has a long vowel followed by the first half of a geminate. The second type has a long vowel closed with the first consonant of a sequence of two consonants; this consonant must be from a very restricted set of consonants, either a coronal or /h/. I am analyzing the first type of syllable as a tri-moraic syllable and the second type as a bi-moraic syllable with an appendix consonant. I will provide a full Optimality Theoretic analysis of this asymmetry, including the relevant constraints, all establishable rankings, and full tableaux for relevant underlying forms. The result of the Finnish case is that a constraint, $*\sigma_{\mu\mu\mu}$, can prove fatal to candidates from one underlying representation while not being appealed to for the evaluation of candidates of another underlying representation. This result is possible in Optimality Theory, but not in approaches where constraints are surface-true, a point returned to in 6.2.

Chapter 4 inspects two languages which are rather similar from our point of view, Fula and Koya. Both languages lack sequences of long vowels and geminates, while both have sequences of long vowels and consonant sequences. In order to have a satisfying theory of what kind of hypercharacterized syllables may occur, we must also have an approach to where they may not occur. Fula and Koya lack the hypercharacterized syllables closed with a geminate found in Finnish. Fula and Koya are also languages with appendix consonants and corresponding hypercharacterized syllables closed with a consonant sequence. These languages argue for the appendix to be an adjoined position to the syllable, because the syllables are word internal (or possibly the final mora of a syllable) instead of a higher position, such as the word, as in Borowsky (1986)[7] or Rubach and Booij (1991). The word-final position may be another version of appendix, found in other languages.

The analysis to Fula presented in Chapter 4 has the advantage over earlier approaches of explaining why the geminate degeminates in order to avoid a tri-moraic syllable (where vowel shortening would have the same effect). Koya is analyzed using a constraint against long vowels, suggested by Paradis (1988) and Prince and Smolensky (1993) and employed in the analysis of Wiyot in chapter 2.

Chapter 5 serves three purposes, the first of which is to explore some languages which have appendixal consonants, including three Australian languages, which have appendixal consonants at the ends of syllables word-medially. Evidence for this analysis is adduced from stress, vowel length, and minimal word constraints. One of the attributes that I am claiming for the appendix consonant is preferences for coronal articulation. Australian languages clearly illustrate of this preference.

The second purpose of Chapter 5 is to explore the structure of the appendix. It is assumed here that the consonant is adjoined directly to the syllable node at the end of the syllable. This representation explains why neither onsets nor appendix consonants are relevant for stress placement or vowel length. However, there are a number of other adjunction sites imaginable in the prosodic hierarchy. Some of these have been proposed in the literature as sites to which appendix consonants are adjoined. The appendix position will be analyzed in terms of Alignment constraints, following McCarthy and Prince (1993b).

The different appendix constraints are derived from a phonetic scale of speed of articulation. Apical coronals are the fastest, followed by non-apical coronals, followed by non-coronals. This phonetic scale provides the ranking of the appendix constraints, via the use of phonetic scales, as in Prince and Smolensky (1993, ch. 5).

Chapter 6 has a section devoted to comparing violable constraints to a more parametric approach. That section will show that the results achieved in earlier chapters are genuinely linked to Optimality Theory. Using the same constraints in parametric fashion does not yield the same results without complications of other elements of the grammar.

The balance of chapter 6 is devoted to questions about representation. Section 6.3 discusses questions about the representation of geminates, particularly representations that employ underlying linking, instead of an underlying mora, to mark geminates in the lexicon.

---

[7] This work adjoins the final consonants to the syllable node, but this adjunction is restricted to the word-final position.

Section 6.4 is devoted to the Coda Condition Approach, which employs restrictions on codas to derive the distribution of consonant sequences. This approach can also embrace the linked representation of geminates taken up in 6.3, so these two sections are closely intertwined. Section 6.5 discusses the use of a the Coda Condition Approach as a theory of hypercharacterization.

Sherer, Tim D. 1994 Prosodic Phonotactics. Available from GLSA Publications, Department of Linguistics, South College, University of Massachusetts, Amherst, MA 01003.

PREFACE TO THE COMPUTER-FILE VERSION

This is the second of six chapters of my dissertation. This version of my dissertation is divided into seven chapters, numbered 0 to 6. Chapters 1 through 6 contain the corresponding chapters of my dissertation. Chapter 0 contains the front matter, as well as the appendix and the sources cited.

This version of my dissertation was taken from the Graduate Linguistics Student Association (GLSA) version, which had several typos corrected over the deposit version. This version differs from the GLSA version with respect to some diagrams. Mainly in tableaux, I have substituted bracketing and subscripted morae for tree diagrams in representing syllables, as below:

{C V_μ} {C V_μ}	=	2 cv syllables
{C $V_{\mu\mu}$}	=	a syllable with a long vowel
{C V_μ C_μ}	=	a bi-moraic closed syllable
{C V_μ C}	=	a mono-moraic closed syllable
{C V_μ { C_μ } V_μ }	=	2 syllables with a geminate consonant

There are still some diagrams in their original forms. I hope that these are recoverable from context.

The pagination is different from the original version. Each chapter is paginated separately. There is a table of contents for each chapter at the beginning of that chapter.

This chapter includes a factorial typology involving several of the constraints introduced in chapter 1. Also, there are analyses of facts from the languages Wiyot and Italian. Another question dealt with here is the content of undelrying form--in particular, the result of specifiying a “weight by position” mora in the lexicon. I also discuss the role of lexical levels in OT.

Tim D. Sherer
TDSherer@Delphi.Com
December, 1994

FACTORIAL TYPOLOGY

CHAPTER 2

FACTORIAL TYPOLOGY

2.1 Introduction

Linguistic theory seeks to explain the formal elements of human language. In doing so, we posit certain elements as being possible elements of a grammar. These possible elements are available to every language learner as part of the innate endowment of Universal Grammar. Although these elements are a fixed set, languages can and do vary from each other in significant ways. The elements of a possible grammar may be combined in different fashions by different grammars.

Optimality Theory offers us a theory of this variation. Optimality Theory is a system of ranked constraints. It is possible for these constraints to be found in different rankings in different languages. Because the number of imaginable rankings is the factorial of the number of constraints under consideration, the process of varying the ranking yields what Prince and Smolensky (1993) call a *factorial typology* (see Prince and Smolensky 1993, especially 3.1 and chapter 6). If our sample of languages were large enough, we would expect to find all rankings of constraints. Of course, this idealization is limited in some ways. First, some constraints are based on phonetic hierarchies or perhaps markedness hierarchies, such as the sonority scale, and these constraints are universally ranked with respect to each other in parallel to the hierarchy they are based on (see Prince and Smolensky 1993 Chapter 5). Because these scale-based constraints are universally ranked with respect to each other, we do not need to consider rankings that contradict these rankings. Second, it is further conceivable that some constraints are universally ranked either by universal grammar or by extra-linguistic factors. Third, as a practical matter, certain constraints do not interact with each other. That is, there are constraints that, in either ranking relationship, will yield the same output as optimal. We do not have to consider their rankings with respect to each other in our investigations, although those constraints might interact with other constraints. Fourth, some constraints might be universal not in their entire form, but in their schema, which are parametric in the values that they take. The actual constraints will have different values in different languages (McCarthy and Prince 1993a, 1993b on alignment). Within these limits, however, differences in constraint ranking provide an elegant approach to the differences between grammars.

Prince and Smolensky (1993, chapter 6) employ the interaction of constraints in different ranking relationships to account for various elements of syllable structure. The theory of the structure of the syllable that they employ is minimal, employing notions that virtually any theory of syllable structure will need to employ. Their results, consequently, are relevant to virtually any theory of syllable structure. This chapter explores a factorial typology with a richer syllable structure.

This chapter is an exercise in laying out the various predictions that follow from permuting several constraints. The constraints discussed are constraints on syllable size and syllable final consonants, with several faithfulness constraints that are necessary for any discussion. As far as I know, these constraints are not subject to any of the conditions on factorial typology discussed above, except that some of them do not interact with each other, so we may proceed with a factorial typology, noting only the cases where constraints do not directly interact. I list the constraints below.

1) Syllable Size Constraint

$*\sigma_{\mu\mu\mu}$ Syllables are maximally bi-moraic.

2) Syllable Final Consonant Constraints

a) $*\mu_{CONS}$ Moraic consonants are not allowed.

b) *APPENDIX Appendix consonants are not allowed.

3) Faithfulness Constraints

a) $PARSE_{SEG}$ Segments must be parsed.

b) $PARSE_{\mu}$ Morae must be parsed.

Because we are considering five constraints, there are 5! or 120 possible rankings. To prevent this permuting of constraints from becoming a daunting task, we will proceed with two or three constraints at a time. Because not all constraints interact with every other constraint, we can break the investigation up into smaller factorial typologies. For expository purposes, we will approach sets of interacting constraints as topics, divided up by the phenomena that they cover. For instance, $PARSE_{\mu}$ and $*\mu_{CONS}$, alone, describe the distribution of geminates after short vowels. Below, I will devote a subsection to those two constraints alone, showing how those two constraints interact to permit or disallow geminates.

This chapter will proceed as follows. I will begin by summarizing the results of Prince and Smolensky (1993, chapter 6) that deal with coda consonants. This will serve as both the base for the factorial typology involving more complicated syllable structure and also as a first glimpse of factorial typology itself. Next, we will turn to moraic consonants, including simple and geminate consonants. The constraint $*\mu_{CONS}$ is relevant to both geminates and moraic simple consonants. That section will show how languages prohibit and permit both simple and geminate moraic consonants and how the two types can co-occur within a language. The next section is devoted to appendixal consonants, including discussion of why we do not expect appendixal consonants to be ambisyllabic in this system. Finally, we turn to hypercharacterized syllables. This topic is naturally linked to the others in that hypercharacterized syllables are always closed syllables and as such they are either closed with moraic or appendixal consonants.

In every section, constraints that are not under investigation are assumed to be high ranking and unviolated. As such, candidates which violate them will be ruled out and will be left out of our discussion. I will make note of this, from time to time, below, to try to prevent any confusion.

2.2 Prince and Smolensky's Syllable Typology

The factorial typology that constitutes the bulk of this chapter is concerned with the possibility of consonant sequences and geminates in languages, as discussed above. This work builds on Prince and Smolensky's factorial typology of syllable structure (Chapter 6 of Prince and Smolensky 1993; chapter 8 is also concerned with syllable structure). In this subsection, I will provide a brief summary of Prince and Smolensky's typology with respect to the existence of codas. This is important because I adopt many of their insights in the rest of the chapter.

Prince and Smolensky begin with the question of how a syllable theory can derive the universal canon of possible syllables, from Jakobson. This is summarized below, adapted from Prince and Smolensky's (113).

4) Universal Syllable Canon

		onsets	
		required	optional
codas	forbidden	CV	(C)V
	optional	CV(C)	(C)V(C)

Prince and Smolensky employ the following four constraints.[1]

5) Constraints Employed by Prince and Smolensky

ONSET	Syllables must have onsets.
NOCODA[2]	Syllables must not have a coda.
PARSE	Segments must be parsed into syllables.
FILL	Syllable positions must be filled.

Because this chapter is concerned with consonant sequences that are a "coda" followed by

[1]I have paraphrased these constraints slightly here to avoid confusion with other parts of this dissertation. None of these changes affects this discussion at all.

[2]called "-COD" in Prince and Smolensky.

an onset, I will focus on Prince and Smolensky's discussion of codas. The ranking of ONSET and NOCODA are not relevant to each other, as Prince and Smolensky show. Either ranking of these two constraints will require that a single intervocalic consonant will be parsed as an onset, not a coda (see the Universal Optimal Syllable Theorem, Prince and Smolensky 1993, section 6.2.2). If we exclude ONSET from discussion, there are six ranking relationships. That is, there are three constraints under discussion and the number of ranking relationships is the factorial of three, which is six. I will discuss a single underlying form for each of these ranking relationships--CVC--which, if parsed faithfully, will have a coda. Prince and Smolensky's chapter is concerned with more issues than just this and I am only summarizing this one element of it (section 6.2.2.2).

The first case to consider is the rankings where NOCODA is undominated. In these cases, there will be no codas. The following tableau has FILL as the lowest constraint.

6) Tableau

	/CVC/	NOCODA	PARSE	FILL
	.CVC.	*!		
☞	.CV.C□.			*
	.CV.<C>		*!	

The result of this constraint ranking is that an epenthetic vowel is predicted to take the second consonant as an onset. Alternatives with a coda and with an unparsed segment are suboptimal.

The next tableau also has NOCODA as undominated.

7) Tableau

	/CVC/	NOCODA	FILL	PARSE
	.CVC.	*!		
	.CV.C□.		*!	
☞	.CV.<C>			*

In this case, the optimal candidate has an unparsed segment. Candidates with a parsed consonant, either as a coda or as the onset to an epenthetic vowel, are sub-optimal.

NOCODA can be dominated in a language without having coda consonants. The next tableau shows a situation where PARSE is undominated and NOCODA dominates FILL.

8) Tableau

	/CVC/	PARSE	NOCODA	FILL
	.CVC.		*!	
☞	.CV.C□.			*
	.CV.<C>	*!		

In this situation, the first candidate, which does have a coda, is still not optimal. Instead, a candidate with an epenthetic vowel will be optimal. This is the same output as the ranking relationship in (6), above.

If FILL, not PARSE, is undominated and NOCODA dominates PARSE, the result will also lack a coda. The following tableau shows this situation.

9) Tableau

	/CVC/	FILL	NOCODA	PARSE
	.CVC.		*!	
	.CV.C□.	*!		
☞	.CV.<C>			*

Here the second consonant is optimally left unparsed. Both candidates where it is parsed, either as a coda or as the onset to an epenthetic vowel, are suboptimal. This is the same result as in (7).

Finally, we turn to cases where both FILL and PARSE dominate NOCODA. The ranking of FILL and PARSE are not relevant here, only the fact that both are above NOCODA. I summarize both situations in the following tableau.

10) Tableau

	/CVC/	PARSE	FILL	NOCODA
☞	.CVC.			*
	.CV.C□.		*!	
	.CV.<C>	*!		

Only when both of these two constraints are ranked above NOCODA will there be a coda. Otherwise, a coda will be sub-optimal.

This approach also predicts that the alternatives to parsing a consonant as a coda are underparsing and overparsing, or epenthesis and deletion, in process-oriented terminology. The problem with process models of phonology, as Prince and Smolensky point out at various points in their work (see especially chapters 3 and 4) is that the different results in these cases are all coming from the same constraint against codas. Process models must have different processes for epenthesis and deletion which have, hidden inside them, the imperative to avoid codas. Prince and Smolensky's approach avoids this missed generalization by stating the generalization that codas are dispreferred and having a mechanism to allow interaction between this constraint and other constraints.

The results in this section were achieved by permuting the ranking of three constraints. The same constraints, in different ranking relationships, predict a range of results. This example of factorial typology shows the benefits of factorial typology. We can, based on constraints proposed for one analysis, make predictions about all grammars.

This chapter is concerned with extending this result to a richer theory of syllable structure. That is, I am concerned here with a factorial typology which employs moraic and appendixal consonants, and makes predictions with respect to simple and geminate moraic consonants, as well as including vowel length.

2.3 Moraic Consonants

Moraic Theory (McCarthy and Prince 1986, 1988, 1990, Hayes 1989, Zec 1988, Itô 1989) deals with two very different types of problems. In addition to its crucial role in the stress systems of languages, where consonants are either relevant or not for the purposes of foot assignment,[3] morae are also important in their role as licensers.[4] This second role is our concern. A mora can be intermediate between a consonant and a syllable. As such, constraints on both a consonant's association to a mora and the constraints on a mora's association to a syllable have a direct effect on the environments in which a consonant may be licensed through a mora and consequently on the possible strings of a language.

Below I give the two types of moraic consonants from Chapter 1.

11) Two Types of Moraic Consonants

MORAIC	CONSONANTS
MORAIC CODA	GEMINATE
σ_1 σ_2	σ_1 σ_2
μ_1 μ_x μ_2	μ_1 μ_x μ_2
V_1 C_1 C_2 V_2	V_1 G V_2

To repeat, a moraic coda is a syllable final consonant with a mora as intermediary to the syllable node and a geminate is an ambisyllabic consonant with a mora as an intermediary to the syllable node.

This section will deal with these structures from a typological point of view. That is, because languages differ on whether or not they have moraic codas and geminates, we want a system that permits variation from one grammar to the next. This system will make predictions about to what extent these different types of moraic consonants show up in different languages and whether the existence of one type entails the existence of the other.

A set of constraints in a particular ranking used to analyze phenomena in a language also implies a typology of possible grammars, as discussed above. Every other ranking is possible, within the limits discussed above, and we expect those rankings to be found in some grammar. The constraints focused on here are given below.

12a) $\text{PARSE}_{\text{SEG}}$ Segments must be parsed.

b) PARSE_{μ} Morae must be parsed.

---

[3]Discussion of this element of Moraic Theory is extensive. In addition to the works noted above, see Hayes (manuscript) and sources cited there for work on stress systems.

[4]Work on this area of moraic theory is also extensive. In addition to the sources above, see Inkelas and Cho 1993, Ito 1989, and Zec 1988, 1993, among others.

c) $*\mu_{CONS}$ Consonants must not be parsed as moraic.

The first two constraints will prohibit the occurrence of stray consonants or unassociated morae, which are not realized at the surface. PARSE constraints are analogous to the prosodic licensing and stray erasure effects of Itô's work (1986). I am adopting a recursive definition of PARSE, which is relevant to any constraint PARSE_X. This definition is given below.

13) Parse (definition):

x is parsed if x is associated to some element higher on
the prosodic hierarchy, y, and y is parsed.

Imagine that a consonant is associated to a mora, which is not parsed, as below.

14)

σ_1			σ_2
μ	$<\mu>$		μ
V_1	C_1	C_2	V_2

With the PARSE definition above, it is clear that the consonant in question, C_1, is not parsed and will not surface. If this is the case, we can use the generalizations that account for the distribution of morae to dictate the distributions of consonant sequences, in combination with other possible ways to parse C_1. Consider the underlying form /vEktor/. If the language in question permits moraic codas, the surface form will be [vEktor], as it is in English. If the language does not permit moraic consonants, the surface form will be [vEtor] (or perhaps the consonant in question will be parsed as the onset to an epenthetic vowel--a FILL violation.) This is the approach taken by most researchers in explaining why a language has consonant sequences--because a consonant can be licensed by a mora at the end of a syllable. Conversely, if no element can PARSE a consonant at the end of a syllable in a language, that language will lack consonant sequences.[5] The third constraint in (12) is a restriction on the occurrence of moraic consonants. This constraint says that moraic consonants are dispreferred. The constraint $*\mu_{CONS}$ is a descendant of NOCODA of Prince and Smolensky (1993), dealing only with moraic codas. A constraint discussed later in this chapter, *APPENDIX, does the other work of NOCODA. The reason that the predicate PARSE is employed in (12c) will be taken up in 4.2.5.

We will proceed by examining how these three constraints interact to predict the two types of moraic consonants. Then we will examine to what extent the two types of moraic consonants are predicted to co-occur within languages by examining the possible rankings relationships between the three constraints.

---

[5] Because Optimality Theory evaluates entire structures, there is no need to question the top-down or bottom-up nature of parsing by recursive definition. PARSE constraints simply check whether something is parsed all the way up.

There are other constraints that could interact with $*\mu_{CONS}$. In particular, the following constraint will turn up in our discussion below, from time to time.

15) $FILL_{\mu}$ Morae must be associated to segmental material.

This constraint militates against any empty morae (see Prince and Smolensky 1993, McCarthy and Prince 1993a). In structures that violate $FILL_{\mu}$, there is an epenthetic vowel. An epenthetic vowel can parse a consonant as the onset to a syllable. Below I repeat a diagram from the introduction, which contains a FILL violation, shown here with "☐".

16) σ_1 σ_x σ_2

μ_1 μ_x μ_2

V_1 C_1 ☐ C_2 V_2

This $FILL_{\mu}$ violation is an alternative to parsing a consonant as moraic, just as leaving a consonant unparsed is an alternative to parsing a consonant as moraic. When either $FILL_{\mu}$ or $PARSE_{SEGMENT}$ is ranked below $*\mu_{CONS}$ in a language, there will not be moraic simple consonants. Either $FILL_{\mu}$ or $PARSE_{SEGMENT}$ will be violated instead. In the following discussion, I refer only to the ranking of $PARSE_{SEGMENT}$, assuming that $FILL_{\mu}$ is unviolated in most cases. I will return to discussion of $FILL_{\mu}$ and $PARSE_{SEGMENT}$ when the language Fijian is discussed in section 2.3.1, below.

2.3.1 Moraic Coda

In line with the general principles of Optimality Theory, when two consonants are adjacent at underlying representation, all possible structures are generated from that lexical representation and those structures are evaluated. When there are two adjacent consonants, one candidate will have moraic coda and another will have an unparsed segment. I repeat the relevant diagrams below.

17) σ_1 σ_2

μ_1 μ_x μ_2

V_1 C_1 C_2 V_2

18) σ_1 σ_2

μ_1 μ_2

V_1 <C_1> C_2 V_2

There are other candidates, of course, such as the one in (16) above, but these are assumed to be ruled out by higher constraints. For this section, we will focus on the alternatives in (17) and (18).

It seems odd, on first blush, to begin discussion of a certain structure with a constraint that prohibits it, but it is the tension between two constraints that sets the limits of variation that we find in grammars (see Prince and Smolensky 1993, especially chapters 8 and 9). One constraint wants to incorporate all the material into the structure, to maintain all of the phonological information from the underlying form. Another wants to make the output conform to certain elements of well-formedness. One or the other constraint is more highly ranked in the language. Prince and Smolensky (1993) refer to this as *push-pull parsing* (chapter 8).

To take a case, let us turn to English. English has moraic consonants. Evidence of this comes from both the stress system and minimal word constraints (McCarthy and Prince 1986; see also Hayes (manuscript) and sources cited there). Consider the underlying form of "garden", given below, with the vowel's underlying morae noted with superscripts.[6]

19) $ga^{\mu}rd\epsilon^{\mu}n$

Below is a tableau showing structures with and without a moraic consonant. These candidates are evaluated by $PARSE_{SEGMENT}$ and $*\mu_{CONS}$.[7] Here we assume that *APPENDIX is high ranked and candidates with appendix consonants are ruled out already.

20) Tableau for $PARSE_{SEGMENT}$ >> $*\mu_{CONS}$

	$/ga^{\mu}rd\epsilon^{\mu}n/$	$PARSE_{SEGMENT}$	$*\mu_{CONS}$
a) ☞	{ g a_μ r_μ }{ d ϵ_μ n }		*
b)	{ g a_μ <r> }{ d ϵ_μ n }	*!	

In this situation, candidate (20a) meets the more highly ranked constraint against unparsed segments, while it violates the less highly ranked constraint against moraic consonants. Candidate (20b) violates only $PARSE_{SEGMENT}$, which is the higher ranking of the constraints. English prefers to parse the consonant as moraic, as opposed to leaving it unparsed, to be unrealized or "deleted". This behavior is in spite of the constraint against moraic consonants. Furthermore, English does not make use of an epenthetic vowel (in violation of $FILL_{\mu}$) which

[6]This notation will be discussed in the section on geminates, below.

[7]The Final /n/ is left out of consideration here.

would also prevent a $*\mu_{CONS}$ violation. The candidate that violates $*\mu_{CONS}$ is the best possible result.

Of course, if a language has the constraints $PARSE_{SEGMENT}$ and $*\mu_{CONS}$ in the opposite ranking, we would expect no consonant sequences beginning with a moraic consonant. Furthermore, this is true of any input whatever. In Optimality Theory, there are no constraints on the input to the constraints, but any input, in the case described here, will be mapped onto an output that lacks consonant sequences.[8] The output of such a language will lack consonant sequences, except, of course, for the possibility of appendixal sequences, not considered here. This is shown in the tableau below, with an example from Fijian (Dixon 1988), a language which has no consonant sequences. Here I give a contrived underlying form, noted in the upper left corner of the tableau, which does have a consonant sequence.

21) Tableau for $*\mu_{CONS}$ >> $PARSE_{SEGMENT}$

	$/ta^{\mu}kta^{\mu}/$	$*\mu_{CONS}$	$PARSE_{SEGMENT}$
a)	{ t a_{μ} k_{μ} } { t a_{μ} }	*!	
b) ☞	{ t a_{μ} <k> } { t a_{μ} }		*

In this tableau, candidate (21a) violates the more highly ranked constraint against moraic consonants, while candidate (21b) meets that constraint and violates the lower ranked $PARSE_{SEGMENT}$. That there are no sequences in Fijian has nothing to do with the underlying representation or input. That fact is a property of the output. Even if there are underlyingly no consonant sequences in Fijian, that is a property not of the lexicon in any theoretic way. Rather, that attribute is a result of the phonology of Fijian providing a surface structure that lacks sequences, which serves as the input for language learners. This line of reasoning is due originally to Stampe (1969, 1973; see Prince and Smolensky 1993, 4.3.1 and also chapter 9).

Fijian is a case where $FILL_{\mu}$ will have the same effect as $PARSE_{SEGMENT}$. Either constraint, when below $*\mu_{CONS}$, will yield a language with no consonant sequences. Consider the following tableau, which has the same contrived underlying form, but with the constraints $FILL_{\mu}$ and $*\mu_{CONS}$.

[8]This line of reasoning is addressed in various places in Prince and Smolensky 1993, not only in the Jakobsonian typology in chapter 6; see especially chapter 4, section 3, also chapter 9, and comments in chapter 10, section 1. Compare this with section 2.3.4.4, below, which deals with learnability.

22) Tableau for $*\mu_{CONS}$ >> $FILL_{\mu}$

	$/ta^{\mu}kta^{\mu}/$	$*\mu_{CONS}$	$FILL_{\mu}$
a)	{ t a_{μ} k_{μ} } { t a_{μ} }	*!	
b) ☞	{ t a_{μ} } { k $□_{\mu}$ } { t a_{μ} }		*

In this tableau, the candidate with an epenthetic vowel is preferred to the candidate with a moraic consonant. In much the same way, the previous tableau shows how interaction with $PARSE_{SEGMENT}$ disprefers moraic consonants. Both outputs have no sequences, consequently both are CV languages. In absence of other evidence, the language learners would not posit an underlying consonant sequence in either case.

In such a language, we can't necessarily tell how the speaker deals with underlying consonant sequences in the input. There are no forms of the language where there are alternations. The roots of Fijian are all vowel final and there are no prefixes. Suffixation does not create underlying consonant sequences. One possible way of determining the question is to inspect loan word phonology. If we assume that the surface from of a loan word in the lending language is taken as a deep structure for the borrowing language (Hyman 1970), we can inspect the surface form of the borrowing language as the output of the phonology. Below is a loan word in Fijian from English.

23) /kis/ "kiss" English

The following tableau shows how this form is mapped to a surface form.

24) Tableau Evaluating loanword "kiss" in Fijian

	$/ki^{\mu}s/$	$*\mu_{CONS}$	$PARSE_{SEG}$	$FILL_{\mu}$
a)	{ k i_{μ} s_{μ} }	*!		
b)	{ k i_{μ} <s> }		*!	
c) ☞	{ k i_{μ} } { s $□_{\mu}$ }			*

Fijian overparses this example. The surface form is [kisi]. Of course, there are more than a few complications in the loan word phonology of Fijian, just as in any language. Fijian seems to employ some parse violations, as well as dealing with sonorant codas different from obstruent codas, at least in some cases. A full analysis will have to be done before drawing any conclusions, which is beyond the scope of this discussion.

To summarize this approach to moraic coda consonants, there are two constraints which are in conflict. $\text{PARSE}_{\text{SEGMENT}}$ and $*\mu_{\text{CONS}}$ are militating against different representations. If $\text{PARSE}_{\text{SEGMENT}}$ is ranked above $*\mu_{\text{CONS}}$, moraic consonants are preferable to unparsed consonants and the language may have consonant sequences which begin with moraic consonants. If $*\mu_{\text{CONS}}$ dominates $\text{PARSE}_{\text{SEGMENT}}$, underparsing the consonant is preferable, leaving no sequences in the language. Similarly, if FILL_{μ} is ranked above $*\mu_{\text{CONS}}$, moraic consonants are preferable to overparsing, and if $*\mu_{\text{CONS}}$ is ranked above FILL_{μ} an overparsing is preferable to a moraic consonant, leaving no consonant sequences in the language.

2.3.2 Geminates

This approach to the representation of geminates in the lexicon follows that of McCarthy and Prince (1986, 1988) and Hayes (1989) in having a geminate be represented in the lexicon as a simple consonant with a mora. Here I draw examples from Woleaian (Sohn 1975), a Trukese language that will serve as an example again below. Below is an underlying representation of a word with and without a geminate.

25)	to be impatient, tired	to be broken
	[basü]	[bünnü]
	/ba$^{\mu}$sü$^{\mu}$/	/bü$^{\mu}$n$^{\mu}$ü$^{\mu}$/

Woleaian has geminates, but lacks consonant sequences, except in very recent loan words from Japanese and English.

The earlier discussion of morae have covered only morae added by GEN. In discussing geminates, we will need to refer to underlying morae, for which I am adopting a notation of superscripting morae in lexical representations, as discussed in 1.3.5. This notation reflects the fact that underlying representations do not contain association lines between the morae and their segmental material. Instead, a mora is co-indexed with underlying featural specification of certain segments, where co-indexing does not entail temporal co-occurrence. The association lines will be provided by GEN, as explained in Chapter 1. To clarify, the representation of a simplex consonant and a geminate consonant is as below, with /p/ and /pp/.

26a) p $[\text{LABIAL}]_i$, $[\text{-SONORANT}]_i$, $[\text{+CONSONANTAL}]_i$, $[\text{-NASAL}]_i$, $[\text{-CONTINUANT}]_i$, ...

b) pp μ_i, $[\text{LABIAL}]_i$, $[\text{-SONORANT}]_i$, $[\text{+CONSONANTAL}]_i$, $[\text{-NASAL}]_i$, $[\text{-CONTINUANT}]_i$, ...

The exact set of features is not important, except that the geminate version is co-indexed with a mora. GEN will take features, like the ones above, as its input and produce all the candidates that are possible from those features, respecting linear order and any existing structure. I include, below, a set of candidates for the two Woleaian examples. First, there are two examples for /ba$^{\mu}$sü$^{\mu}$/.

27a) σ_1 σ_2

μ μ

b a s ü

b) σ_1 σ_2

μ μ μ

b a s ü

The example in (27a) has the /s/ parsed as an onset. See Prince and Smolensky (1993) for an explanation for why it is not ever parsed exclusively as a coda (Prince and Smolensky 1993 Chapter 6). Example (27b) has /s/ as not only parsed as an onset, but also associated to a mora. The mora is provided by GEN. Both of these two forms will be candidates and our system should select (27a) as the optimal form for this input in Woleaian. The form in (27b), however, is found in languages as the optimal output for inputs like the one being discussed. See section 2.7.1 for discussion of forms like these.

Below are examples for the underlying form /bü$^{\mu}$n$^{\mu}$ü$^{\mu}$/.

28a) σ_1 σ_2

μ μ μ

b ü n ü

b) σ_1 σ_2

μ <μ> μ

b ü n ü

The candidate in (28a) has a mora that dominates the /n/. The mora is parsed and has content, but by being associated to a consonant violates $*\mu_{CONS}$. The candidate in (28b) has an unparsed mora and therefore violates PARSE$_{\mu}$, but will not violate $*\mu_{CONS}$. In other words, the two representations differ in that each violates one of two constraints and meets the other. Here, again, we have a conflict between two elements of a grammar. On the one hand, morae must be preserved, on the other, moraic consonants are illicit.

There are a couple of other forms that should be discussed at this point.

29a) σ_1 σ_2

$\mu \langle\mu\rangle$ μ

b ü n ü

b) σ_1 σ_2

μ μ μ

b ü n ü

The first case has an unparsed mora which is not associated to the /n/. Recall the exact statement of $*\mu_{CONS}$. Linking the mora to the segment to which it is co-indexed will not affect that constraint. See 4.2.5 for discussion of Fula, where there is evidence that there is a link between a segment and an unparsed mora.

The representation in (29b) involves the association of the mora to the preceding vowel melody. This is the representation of a compensatory lengthening. This candidate may be optimal in some languages. In the case of languages with geminates, the constraint Segmental Integrity, as discussed in 1.3.5, must be undominated.

To return to the cases in (28), the two constraints of interest here are $*\mu_{CONS}$ and $PARSE_\mu$. To see how these two constraints interact we will examine the actual examples for Woleaian in a tableau, below. The constraints are ranked, here, with $PARSE_\mu$ above $*\mu_{CONS}$.

30) Tableau for $PARSE_\mu >> *\mu_{CONS}$

	$/bü^\mu n^\mu ü^\mu/$	$PARSE_\mu$	$*\mu_{CONS}$
a) ☞	$\{ b\ ü_\mu \{ n_\mu \}\ ü_\mu \}$		*
b)	$\{ b\ ü_\mu \} \{ n_{\langle\mu\rangle}\ ü_\mu \}$	*!	

In this situation, candidate (30a) meets the more highly ranked constraint against unparsed morae, while it violates the less highly ranked constraint against moraic consonants, while candidate (30b) violates $PARSE_\mu$. Woleaian parses the geminate's mora, despite the constraint against moraic consonants. The constraint *μ_C is not active in this case because any candidate where it is not violated has already been removed from consideration by the higher ranking PARSE$_\mu$ (see Prince and Smolensky 1993, section 5.3 on Panini's Theorem.)

Next, we need to see that other consonants will not become geminates willy nilly. Consider the other underlying form given above, /ba$^\mu$sü$^\mu$/, which does not surface with a geminate, in a tableau with the same ranking as above.

31) Tableau for PARSE$_\mu$ >> *μ_{CONS} without Geminate

	/ba$^\mu$sü$^\mu$/	PARSE$_\mu$	*μ_{CONS}
a) ☞	{ b a$_\mu$ } { s ü$_\mu$ }		
b)	{ b a$_\mu$ { s$_\mu$ } ü$_\mu$ }		*!

In this tableau, we see that the first candidate does not violate either of these two constraints. The second candidate does violate a constraint. It violates *μ_{CONS}. Unlike the previous example, the higher ranked constraint, PARSE$_\mu$, is not violated by the other candidate, (31a), because there is no mora in underlying representation. Therefore, the candidate that violates *μ_{CONS} is not optimal.

The two underlying representations discussed here, /ba$^\mu$sü$^\mu$/ and /bü$^\mu$n$^\mu$ü$^\mu$/, have different optimal outputs under this constraint ranking. In (30a), the mora must be parsed because PARSE$_\mu$ dominates *μ_{CONS}, while in (31a), there is no mora to be parsed. The property of Containment is vital to this point. Example (30) lacks a candidate like (31a) with no third mora because Containment requires all the underlying morae in /bü$^\mu$n$^\mu$ü$^\mu$/ to be part of every candidate, even if unparsed and unrealized.

We can see that a language with the opposite ranking of these two constraints will not have geminates. We need to look at two cases, just as we did in Woleaian. One case has a moraic consonant in underlying representation and one case has no moraic consonant. There are plenty of languages like this, Fijian, for example, or English. I give schematic examples here. The following tableau shows an example with an underlyingly moraic consonant, "G", for "geminate".

32) Tableau for $*\mu_{CONS}$ >> $PARSE_{\mu}$

	$/V^{\mu}G^{\mu}V^{\mu}/$	$*\mu_{CONS}$	$PARSE_{\mu}$
a)	$\{ V_{\mu} \{G_{\mu} \} V_{\mu} \}$	*!	
b) ☞	$\{ V_{\mu} \} \{ G_{<\mu>} V_{\mu} \}$		*

In this tableau, candidate (32a) violates the more highly ranked constraint against moraic consonants, while candidate (32b) meets that constraint and violates the lower ranked $PARSE_{\mu}$. Such a language maps underlying geminates to surface as simple consonants.

It does not matter whether the mora is underlying or not. Consider the following case where the mora is added by GEN. In this tableau, we see that even though there is a candidate where there is a mora supplied by GEN, that candidate will be worse than the candidate where the consonant is parsed only as an onset. The candidate with a mora inserted by GEN will violate $*\mu_{CONS}$.

33) Tableau for $*\mu_{CONS}$ >> $PARSE_{\mu}$ without Geminate

	$/V^{\mu}CV^{\mu}/$	$*\mu_{CONS}$	$PARSE_{\mu}$
a)	$\{ V_{\mu} \{ C_{\mu} \} V_{\mu} \}$	*!	
b) ☞	$\{ V_{\mu} \} \{ C V_{\mu} \}$		

In this section, the interaction of two constraints has been shown to give rise to geminates. One constraint is $PARSE_{\mu}$ and the other is $*\mu_{CONS}$. $PARSE_{\mu}$ requires that a mora, which is marked underlyingly on the segment, must be parsed. Parsing it, however, gives rise to a violation of $*\mu_{CONS}$. If $PARSE_{\mu}$ is the higher, then morae must be parsed and the language may have geminates. If $*\mu_{CONS}$ is higher then $PARSE_{\mu}$ is violated and the language will never have geminates.

2.3.3 Consonant Sequences and Geminates

Both geminates and consonant sequences arise as the result of the interaction of constraints. Both arise when $*\mu_{CONS}$ is subordinated to another constraint, $PARSE_{\mu}$ or $PARSE_{SEGMENT}$. This section is devoted to a typology which includes geminates and consonant sequences in the same languages.

The existence of geminates and consonant sequences in languages cross-classifies. That is, there are languages with consonant sequences and no geminates, and there are languages with geminates and no consonant sequences. There are also languages with both geminates and consonant sequences and neither consonant sequences nor geminates. These alternatives are summarized below.

34) Relationship of Sequences and Geminates

	with sequences	without sequences
with geminates	Japanese, Arabic	Woleaian, Ojibwa
without geminates	English, Gumbaynggir	Fijian

Again, we are concerned here with the prosodic effects. Languages can and do disallow some sequences, even when they have others of the same type. Japanese and Arabic both have sequences that begin with moraic consonants. Arabic, though, allows many more sequences than Japanese.

What permits this cross-classification is the property of Containment. A geminate, at underlying representation, has a mora. That mora must be part of every candidate, even if it is not parsed. A consonant sequence has no geminate in the lexicon. An underlying representation with a consonant sequence will have a mora associate to a consonant in some candidates and lack such a mora entirely in other candidates.

We can make the prosodic predictions in the table above using the three constraints discussed so far. These are repeated below.

35a) PARSE_{μ} Morae must be parsed.

b) $*\mu_{\text{CONS}}$ Moraic consonants are not allowed.

c) $\text{PARSE}_{\text{SEGMENT}}$ Segments must be parsed.

To see how the interaction of the constraints works to produce a typology, we must examine the different possible rankings of these three constraints, as a factorial typology. With these constraints, there are six possible ranking relationships (that is, 3 X 2 X 1), but PARSE_{μ} and $\text{PARSE}_{\text{SEG}}$ do not directly interact. Leaving a mora unparsed in a candidate will never allow a segment to be parsed in that form and leaving a segment unparsed will never allow a mora to be parsed. Below, I include diagrams of a form with an unparsed mora and an unparsed segment to illustrate this.

36a) σ b) σ

μ <μ> μ

V C V <C>

This non-interaction reduces the distinct ranking relationships to four. I list these below.

37a) $PARSE_{\mu}$, $PARSE_{SEGMENT}$ >> $*\mu_{CONS}$

b) $*\mu_{CONS}$ >> $PARSE_{\mu}$, $PARSE_{SEGMENT}$

c) $PARSE_{\mu}$ >> $*\mu_{CONS}$ >> $PARSE_{SEGMENT}$

d) $PARSE_{SEGMENT}$ >> $*\mu_{CONS}$ >> $PARSE_{\mu}$

In (37a) and (37d), $PARSE_{SEGMENT}$ dominates $*\mu_{CONS}$, so consonant sequences may exist. In (37b) and (37c), $*\mu_{CONS}$ dominates $PARSE_{SEGMENT}$, so there can be no consonant sequences. In (37a) and (37c), $PARSE_{\mu}$ dominates $*\mu_{CONS}$, so there can be geminates. In (37b) and (37d), $*\mu_{CONS}$ dominates $PARSE_{\mu}$, so there will be no geminates. Consequently, (37a) permits both sequences and geminates, (37b) permits neither, (37c) permits only geminates, and (37d) permits only consonant sequences.

Below, we take these four distinct cases one at a time, with brief tableaux for example languages for each case. The figures below are actually two tableaux per example, with a double line separating the geminate candidates from the consonant sequence candidates. Our goal is to derive the chart above, where geminates and consonant sequences can exist in the same language, but where neither is implied by the other. Once again, we leave appendixal consonants out of our considerations.

We begin our typology with a language where there are both consonant sequences and geminates, as in (37a). Our example is Japanese.[9] Below are a tableaux for these two forms.

38a) /$ni^{\mu}s^{\mu}i^{\mu}$/ b) /$ra^{\mu}mbo^{\mu\mu}$/
[nis:i] [rambo:]
"diary" "rough"

[9]For another approach to Japanese, see Ito 1986, 1989. Her approach makes use of linked structures to predict that Japanese will have only geminates and linked clusters, predictions that the current system does not make. The use of linking is discussed in Chapter 6.4, below.

39) Tableaux for Geminates and Sequences in Japanese

		$PARSE_\mu$	$PARSE_{SEGMENT}$	$*\mu_{CONS}$
a) ☞	$\{ n\ i_\mu \{ s_\mu \} i_\mu \}$			*
b)	$\{ n\ i_\mu \} \{ s_{<\mu>}\ i_\mu \}$	*!		
c) ☞	$\{ r\ a_\mu\ m_\mu \} \{ b\ o_{\mu\mu} \}$			*
d)	$\{ r\ a_\mu <m> \} \{ b\ o_{\mu\mu} \}$		*!	

Japanese has $*\mu_{CONS}$ below both of the PARSE constraints. Consequently, Japanese parses the geminate's mora and parses the first consonant of the sequence as a moraic consonant. This arrangement of constraints predicts both geminates and consonant sequences. Another candidate is possible for the underlying form with a consonant sequence, given below.

40) Form with unparsed mora and segment

```
   σ              σ

   μ <μ>          μ μ

 r a <m>     b    o
```

Under the definition of PARSE adopted here, this form will still violate $PARSE_{SEGMENT}$, in addition to $PARSE_\mu$. That is, the /m/ will not be considered as parsed since it is not attached to a parsed element. Such a candidate will always incur one extra violation over a candidate like (39d) and therefore be less harmonic than that candidate.

Next, we'll consider a language that has neither consonant sequences nor geminates, in this case, Fijian. This tableau has $*\mu_{CONS}$ above both of the PARSE constraints, as in (37b), instead of below them, as in (37a). These examples are all contrived to show that Fijian would not permit such underlying representations to be faithfully parsed. That is, Fijian does not have such representations, but it is a goal of Optimality Theory (see Prince and Smolensky 1993, chapters 2 and 3) that only the surface constraints are of importance, not the underlying forms. Generalizations are not captured through means of lexical restrictions, which are used to rule out certain forms within the lexicon, only to have that fact repeated later in the derivational component of the language.

41) Geminates and Sequences in Fijian

		$*\mu_{CONS}$	$PARSE_{\mu}$	$PARSE_{SEG}$
a)	{ b a_{μ} { s_{μ} } u_{μ} }	*!		
b) ☞	{ b a_{μ} } { $s_{<\mu>}$ u_{μ} }		*	
c)	{ b a_{μ} t_{μ} } { k u_{μ} }	*!		
d) ☞	{ b a_{μ} <t> } { k u_{μ} }			*

In Fijian, or another CV language,[10] it is more harmonic to leave the geminate's mora unparsed and to leave the first consonant of a sequence unparsed than to parse them. Consequently, neither geminates nor sequences will exist on the surface.

The existence of geminates and sequences are completely independent in this system. A language may have both or neither, as above, where $*\mu_{CONS}$ is below or above both PARSE constraints. When $*\mu_{CONS}$ intervenes between the two parse constraints, the language will have either geminates or consonant sequences, but not both.

One such language is Woleaian, which was discussed above. Woleaian has only geminates. In Woleaian, $PARSE_{\mu}$ is above $*\mu_{CONS}$, which is in turn above $PARSE_{SEGMENT}$, as in (37c). The following tableau repeats the Woleaian geminate example, as well as a contrived underlying representation with a consonant sequence.

[10]Recall the discussion of Fijian, in 2.3.1, where $FILL_{\mu}$ has the same result.

42) Geminates and Sequences in Woleaian

		PARSE$_\mu$	$*\mu_{CONS}$	PARSE$_{SEG}$
a) ☞	{ b ü$_\mu$ { n$_\mu$ } ü$_\mu$ }		*	
b)	{ b ü$_\mu$ } { n$_{<\mu>}$ ü$_\mu$ }	*!		
c)	{ b ü$_\mu$ n$_\mu$ } { t ü$_\mu$ }		*!	
d) ☞	{ b ü$_\mu$ <n> } { t ü$_\mu$ }			*

In this tableau, it is worse to leave the geminate's mora unparsed than it is to parse it. Nevertheless, it is worse to have a moraic consonant as the first member of a sequence than it is to leave a consonant unparsed. In other words, even though the language has moraic consonants, those consonants will be restricted to geminates.

If the Woleaian ranking is completely reversed, we get the opposite output, with no geminates, but with consonant sequences, as below. This is the ranking in (37d). We return to English, which has no geminates, but has sequences. This tableau has a contrived underlying moraic consonant and the example "garden" from earlier in this chapter.

43) Geminates and Sequences in English

		PARSE$_{SEG}$	$*\mu_{CONS}$	PARSE$_\mu$
a)	{ a$_\mu$ { t$_\mu$ } ϵ_μ }		*!	
b) ☞	{ a$_\mu$ } { t$_{<\mu>}$ ϵ_μ }			*
c) ☞	{ g a$_\mu$ r$_\mu$ } { d ϵ_μ n }		*	
d)	{ g a$_\mu$ <r>} { d ϵ_μ n}	*!		

Here it is better to violate the constraint against unparsed morae than it is to have a moraic consonant. It is still better to have a moraic consonant as part of a sequence than to have an unparsed segment. Consequently, sequences, but not geminates, are predicted.

This complete cross-classification is possible because geminates and sequences "get" their morae from different sources. The geminate has a mora in its lexical representation and leaving that mora unparsed gives rise to a violation. Leaving a simple consonant unparsed gives rise to a different violation, since a geminate with an unparsed mora is still a parsed segment, by virtue of the onset.

This approach makes the prediction that languages may have geminates, sequences, both, or neither. It makes this prediction using only the different ranking of constraints. Three constraints are discussed here, PARSE_{μ}, $\text{PARSE}_{\text{SEGMENT}}$, and $*\mu_{\text{CONS}}$, which interact to produce this result.

This system does not speak to the nature of the sequences or geminates in question, dealing as it does only with the prosodic elements. Languages can and do have restrictions on the sequences that they have (Prince 1984, Yip 1989, 1991, Lamontagne 1993). The approach of this dissertation handles these only in connection with APPENDIX constraints (see 5.2 and chapter 6).

Japanese and Arabic make a good comparison here. In this system Japanese and Arabic have identical rankings of these three constraints: PARSE_{μ}, $\text{PARSE}_{\text{SEG}} >> *\mu_{\text{CONS}}$. Nevertheless, Japanese sequences are very restricted; only nasal-stop homorganic sequences are allowed. Arabic has a richer set of sequences. Other theories of gemination and theories of feature geometry make predictions about both the relationship of geminates and consonant sequences and about the possible consonant sequences of a language. In discussing these different predictions, we must disentangle the parts of an analysis that follow from representations and those that follow from the theory of Optimality Theory. These questions will be addressed in chapter 6, especially 6.3.

2.4 Appendixes

The appendix is the other means of licensing a consonant at the end of a syllable. The appendix plays a role in various theories of the syllable, as discussed in the introduction. I define an appendix as a consonant that is dominated directly by the syllable node at the end of the syllable. Certain consonants will occur on the surface where they are neither onsets nor moraic. This first consonant of a consonant sequence, in some languages, will be inert with respect to vowel length and with respect to stress. These consonants are what I am referring to as appendix consonants. This inertness suggests that these consonants have no representation on the moraic tier. Furthermore, some languages require consonants at the end of a syllable to be restricted to a single place of articulation (always coronal, often a subset of coronals) (Paradis and Prunet 1992). I am suggesting here that these asymmetries are a fact about non-moraic syllable final consonants.

Below I give a diagram showing the form of an appendix.

44) Appendix: σ

$C]_\sigma$

An appendix is an adjoined position of the syllable.[11] As such, two properties will follow directly from the representation proposed. First, appendixes will not affect stress or other elements of prosody, because the consonant is not dominated by a mora. Second, the appendix will not effect vowel length, again because the consonant is not dominated by a mora.

I am assuming that appendix consonants and moraic consonants are "phonetically identical." The reason for this assumption is that I know of no language where it is claimed that a difference can be heard between moraic and appendixal consonants. Such a difference is certainly never enough to maintain a minimal pair. There may be measurable differences between the two types of consonants, but in absence of data, I will assume that the two are indistinguishable phonetically.

Parsing consonants as appendixes is analyzed in push-pull fashion. There is a constraint against appendixes, in opposition to faithfulness constraints, like $\text{PARSE}_{\text{SEGMENT}}$ and FILL_μ.

45) *APPENDIX: *σ

$C]_\sigma$

The constraint against appendix consonants is parallel to the constraint against moraic consonants. Together, they do the work of NOCODA in Prince and Smolensky's work, although Prince and Smolensky (1993) do not deal with appendixes per se (except in passing in section 4.5; see also McCarthy and Prince 1993b, p. 17-18, where appendixes are analyzed as following from Alignment constraints). The nature of appendix consonants will be taken up in chapter 5, below.

From the point of view of phonotactics, the appendix position adds to the range of possible sequences. A language may rule out consonant sequences that begin with moraic consonants but allow sequences that begin with an appendix consonant.

We would not expect an appendix consonant to be ambisyllabic and this is one way in which they differ from geminates. When a consonant is parsed as appendixal, it is in violation of the constraint *APPENDIX. Consonants are parsed as appendixal to satisfy $\text{PARSE}_{\text{SEG}}$, but if such a consonant is already parsed as an onset, $\text{PARSE}_{\text{SEG}}$ satisfied. This is shown in a tableau

---

[11]We may want to consider similar adjoined positions to the foot or to the word, for languages like English (Borowsky 1986) which have only a word final appendix not a syllable final one (see Rubach and Booij 1991). This taken up at greater length in Chapter 5, especially sections 5.4 and 5.5.

below, with a schematic example. This example shows three candidates. Candidate (46a) has an ambisyllabic consonant. Candidate (46b) has the consonant parsed exclusively as an onset. Candidate (46c) has the consonant parsed exclusively as an appendix consonant. Of course, in this case, we are assuming that the constraint against moraic consonants is undominated and, consequently, we do not have to discuss any candidate that has a moraic consonant, either a consonant parsed exclusively as moraic or a consonant which is ambisyllabic with a mora, that is, a geminate.

46) Schematic Example of Intervocalic Consonant

	/CV$^{\mu}$CV$^{\mu}$/	ONSET	*APPENDIX
a)	{ C V$_{\mu}$ C } { V$_{\mu}$ }	*	*
b) ☞	{ C V$_{\mu}$ } { C V$_{\mu}$ }		
c)	{ C V$_{\mu}$ { C } V$_{\mu}$ }		*

This tableau shows that a single intervocalic consonant will always be parsed as an onset, following Prince and Smolensky (1993, chapter 6). While a candidate parallel to (46c) might be selected for an underlyingly moraic consonant, that will only happen to meet the constraint PARSE$_{\mu}$. There is no candidate parallel to that for appendix consonants. Neither is there room in the system for an underlying appendixal consonant. That would involve a syllable being incorporated into underlying representation, a move not warranted by any data that I am aware of.

The constraint *APPENDIX interacts with both PARSE$_{SEGMENT}$ and *μ_{CONS}, as well as FILL$_{\mu}$, which we are still leaving out of the main of this discussion. If *APPENDIX is ranked above either of these two constraints, it makes no difference. A candidate which has an appendix (such candidates have not been evaluated in the section on moraic consonants, above) will always be less harmonic than candidates which parse the consonant as moraic or fail to parse it, as the case may be. Below I provide a tableau for a contrived representation. This is the ranking suggested in 2.3.3, above for Fijian.

47) Tableau for Unparsed Consonant

	$/ta^{\mu}kta^{\mu}/$	*APPENDIX	$*\mu_{CONS}$	$PARSE_{SEG}$
a)	{ t a_{μ} k_{μ} } {t a_{μ} }		*	
b) ☞	{ t a_{μ} <k> } {t a_{μ} }			*
c)	{ t a_{μ} k } {t a_{μ} }	*!		

Our discussion of *APPENDIX from the point of view of factorial typology will be minimal, since only when the constraint against appendixes is the lowest of the relevant constraints will a language have appendix consonants. The term "relevant" here means that the constraints in question will separate[12] the candidates with respect to the existence and prosodic structure of consonant sequences. A case where *APPENDIX is the lowest constraint is shown in the tableau below. This example is from Koya, which is discussed at greater length in Chapter 4.

48) Tableau for Appendixal Parse in Koya

	/banda/	$PARSE_{SEG}$	$*\mu_{CONS}$	*APPENDIX
a)	{ b a_{μ} n_{μ} } {d a_{μ} }		*!	
b)	{ b a_{μ} <n> } {d a_{μ} }	*!		
c) ☞	{ b a_{μ} n } {d a_{μ} }			*

In this case, the parse with the appendix is preferred. Both a parse with the segment underparsed, as in (48b), and with a moraic consonant, as in (48a), are ruled out by the higher ranking constraints $PARSE_{SEGMENT}$ and $*\mu_{CONS}$. The ranking of $PARSE_{SEGMENT}$ and $*\mu_{CONS}$ is not

[12]See Prince and Smolensky 1993, chapter 5 for a complete definition of the term "separate".

establishable based on this form, but can be established for Koya (see 4.3.3.2). Of course, the candidates in (48a) and (48c) are phonetically identical, meaning that we have to judge between them by other evidence, such as vowel length alternations or Iambic underparsing in Koya (see Chapter 4) or by appendix place asymmetries, where appendixes favor coronality, as in Finnish (see Chapter 3).

The appendix is an option for parsing a consonant. When *APPENDIX is ranked below both $*\mu_{CONS}$ and $PARSE_{SEG}$, a consonant will be parsed as appendixal. Otherwise, it will be parsed as moraic or left unparsed. Appendix consonants are phonetically identical to moraic simple consonants. Consequently, arguing that a consonant is appendixal must involve other evidence beyond the mere existence of a sequence, such as evidence from vowel length alternations, foot assignment or adjustment, and appendix asymmetries of articulator.

2.5 Hypercharacterized Syllables

As defined in Chapter 1, a hypercharacterized syllable is a syllable with a long vowel and a closing consonant or with a short vowel and two closing consonants or any syllable with a short or long vowel and more than two closing consonants. There are several kinds of syllables that meet this definition. One variety has three morae, including two vowel morae and one consonant mora. Such syllables may be closed by a simple moraic consonant or by a geminate. Another type of hypercharacterized syllable has either one or two vowel morae and one or more appendixal consonants.

We will turn to tri-moraic syllables first. In considering these syllables from the point of view of factorial typology, we need to consider the constraints that give rise to moraic simple consonants and to geminates and the constraint $*\sigma_{\mu\mu\mu}$, which prohibits tri-moraic syllables. Again, it is the tension between the constraint that prohibits a structure and the constraints that wish to faithfully represent underlying phonological information that gives our understanding of the possibility of tri-moraic syllables.

Hypercharacterized syllables closed with an appendix are not militated against by a special constraint on a par with $*\sigma_{\mu\mu\mu}$. That is, they are disfavored by *APPENDIX, just as hypercharacterized syllables with a moraic consonant are in violation of $*\mu_{CONS}$. However, hyperchacterized syllables closed with a moraic consonant also violate $*\sigma_{\mu\mu\mu}$, while hypercharacterized syllables closed with an appendix violate no analogous constraint. This system we makes the prediction that vowel length is irrelevant to appendixes. Consequently, a language with appendix consonants should have them after both long and short vowels.

2.5.1 Tri-moraic Hypercharacterized Syllables

The constraint $*\sigma_{\mu\mu\mu}$ was introduced briefly in chapter 1. This constraint is violated by a syllable with three or more morae.

49) $*\sigma_{\mu\mu\mu}$ Syllables are maximally bi-moraic.

A number of researchers have made use of this constraint in one form or another, often to explain the cause of closed syllable shortenings in various languages (see, for instance, Clements 1986 on *V-trimming* in LuGanda and Myers (1985, 1987) on vowel length in English). As a universal constraint, it has been suggested by Lowenstamm and Kaye (1985), McCarthy and Prince (1986), Lowenstamm (1987), Steriade (1991), Paradis (1992 (a translation of a 1986 work)).

The view of this constraint has always been in the nature of a parameter. A language either obeys it or not. The contribution of Optimality Theory is that this constraint, like all others, can now be viewed as a part of all grammars. This constraint, like all others, is violated in certain forms. Here a higher ranking constraint is met and not violated in other forms, even at the expense of lower ranking constraints.

Broselow (1992) proposes that a bi-moraic syllable is optimal in her analysis of epenthesis in several dialects of Arabic. This was not an Optimality Theoretic approach, but let us consider the ramifications of this move in Optimality Theory. Perhaps this approach explains why tri-moraic syllables are less preferred than non-tri-moraic syllables. Suppose we were to adopt the following constraint, as in 2.5.1. This constraint is very reminiscent of Foot-Binarity (Prince 1980, McCarthy and Prince 1986, 1993a, 1993b, Prince and Smolensky 1993, 4.3.1) except that Foot Binarity is only a lower limit, not an upper and lower bound.

50) $\sigma_{\mu\mu}$ Bi-moraic syllables are optimal.

Because constraints are violable, we know that this constraint would not force us to have only bi-moraic syllables. As shown below, tri-moraic syllables, in violation of $\sigma_{\mu\mu}$, can arise to meet PARSE constraints. Underlyingly mono-moraic vowels could be incorporated into bi-moraic syllables, but only at a cost. Adding a mora is possible for GEN, but that is adding a structural position, just as adding a whole syllable is adding a structural position. If there were a cost for adding a mora, in the form of a constraint violation, we can see how mono-moraic syllables could arise. On the other hand, $\sigma_{\mu\mu}$ could be high-ranking, as it is in Wiyot, requiring virtually all syllables to be bi-moraic.

There is a problem with the use of $\sigma_{\mu\mu}$ to replace $*\sigma_{\mu\mu\mu}$. That problem is that $\sigma_{\mu\mu}$, in the right ranking relationship with ONSET and NOCODA, predicts that the underlying form /...VCV.../ could be mapped to [...VC.V...], where /C/ is parsed as a coda, not an onset. This contradicts the Jakobsonian typology (this was discussed in 3.2, above, and in Prince and Smolensky 1993, chapter 4). Although I can't see a way around this prediction, perhaps it does not follow for some reason that I'm not considering.

Another possible way to understand $*\sigma_{\mu\mu\mu}$ comes from Selkirk (1992). Selkirk proposes that syllable structure be derived form parametric principles that are independently motivated. McCarthy and Prince (1986) make a similar point, from the point of view of locality and the inability of grammars to count. In particular, if the morae of a syllable are required to be adjacent to each other, a tri-moraic syllable cannot meet that requirement. On this view, the two mora limit is a structural constraint, now more like *APPENDIX or ONSET.

Whatever the actual form of the constraint is, tri-moraic hypercharacterized syllables are, by their very nature, situations of conflict. They always give rise to a violation of the constraint $*\sigma_{\mu\mu\mu}$. A language may have $*\sigma_{\mu\mu\mu}$ as a highly ranked constraint and not have any hypercharacterized syllables as a result. Candidates which violate $*\sigma_{\mu\mu\mu}$ would be ruled out. Other candidates, which include PARSE violations or FILL violations would be preferable. If the relevant PARSE or FILL constraints are more highly ranked than $*\sigma_{\mu\mu\mu}$, then tri-moraic syllables could arise in some candidates. This sort of conflict is the very stuff of Optimality Theory.

2.5.1.1 Tri-moraic Syllables Closed with a Simple Consonant

A factorial typology of tri-moraic hypercharacterized syllables closed by a simple consonant which is moraic involves the following constraints: PARSE_{μ}, $\text{PARSE}_{\text{SEG}}$, $*\sigma_{\mu\mu\mu}$, $*\mu_{\text{CONS}}$. These four constraints are concerned with syllable size, moraic content, and faithfulness of both segments and morae. In order to have hypercharacterized syllables, the following three rankings must exist.

51a)	$\text{PARSE}_{\text{SEG}}$	>>	$*\mu_{\text{CONS}}$
b)	PARSE_{μ}	>>	$*\sigma_{\mu\mu\mu}$
c)	$\text{PARSE}_{\text{SEG}}$	>>	$*\sigma_{\mu\mu\mu}$

Ranking (51a) is necessary for the language to have moraic simple consonants, whether after a long vowel or after a short. This has been discussed above, in section 2.3.1.

The constraint PARSE_{μ} must dominate $*\sigma_{\mu\mu\mu}$ in order for both of the vowel morae to be parsed. Below is a schematic example with these two constraints. In the following tableau, we see that example (52a) is parsed in violation of $*\sigma_{\mu\mu\mu}$. Candidate (52b) meets $*\sigma_{\mu\mu\mu}$ only at the expense of violating the higher ranking PARSE_{μ}. Consequently, $*\sigma_{\mu\mu\mu}$ must be below PARSE_{μ} in order for this sort of hypercharacterized syllable to arise. Note that this is similar to the ranking that gives rise to hypercharacterized syllables closed with a geminate.

52) Tableau for PARSE_{μ} >> $*\sigma_{\mu\mu\mu}$

	$/V^{\mu\mu}CCV^{\mu}/$	PARSE_{μ}	$*\sigma_{\mu\mu\mu}$
a) ☞	{ $V_{\mu\mu}$ C_{μ} } { C V_{μ} }		*
b)	{ $V_{\mu\langle\mu\rangle}$ C_{μ} } { C V_{μ} }	*!	

If the ranking above were reversed, the result would be an underparsing of a vocalic mora--what we would call closed syllable shortening in a process oriented system. For an example, consider English, which has a few cases of such underparsings. These were discussed in Prince and Smolensky (1993, section 10.3.3.1) following Myers (1987).

53) Tableau for English Vowel Mora Underparsing

	/ki$^{\mu\mu}$pt/	*$\sigma_{\mu\mu\mu}$	PARSE$_{\mu}$
a)	{ k $i_{\mu\mu}$ p_{μ} } t	*!	
b) ☞	{ k $i_{\mu<\mu>}$ p_{μ} } t		*

In this tableau, we have the underlying form with a long vowel and a moraic consonant. The final /t/ is assumed to be appendixal. Candidate (53a) is a faithful
parse of all underlying material, while (53b) leaves unparsed a mora from the vowel /i/. This is the preferred case from this ranking and the actual output form.

The third ranking, PARSE$_{SEG}$ >> *$\sigma_{\mu\mu\mu}$, is necessary for both consonants to be parsed. Below, I give a schematic example of this ranking.

54) Tableau for PARSE$_{SEG}$ >> *$\sigma_{\mu\mu\mu}$

	/V$^{\mu\mu}$CCV/	PARSE$_{SEG}$	*$\sigma_{\mu\mu\mu}$
a) ☞	{ $V_{\mu\mu}$ C_{μ} } { C V_{μ} }		*
b)	{ $V_{\mu\mu}$ <C> } { C V_{μ} }	*!	

In this tableau, we see that example (54a) is parsed in violation of *$\sigma_{\mu\mu\mu}$. Candidate (54b) meets *$\sigma_{\mu\mu\mu}$ by leaving a consonant unparsed, at the expense of violating the more highly ranking PARSE$_{SEG}$. Consequently, *$\sigma_{\mu\mu\mu}$ must be below PARSE$_{\mu}$ in order for this sort of tri-moraic hypercharacterized syllable to arise.

If this situation were reversed, we would have the underparsing of a consonant. I analyze Finnish along these lines. Below, I give an example from Finnish, where /i$^{\mu\mu}$mpi/, a contrived underlying form, is mapped to [iipi].

55) Tableau for Segment Underparsing in Finnish

	/$i^{\mu\mu}mpi^{\mu}$/	*$\sigma_{\mu\mu\mu}$	$PARSE_{SEG}$
a)	{ $i_{\mu\mu}$ m_{μ} } { p i_{μ} }	*!	
b) ☞	{ $i_{\mu\mu}$ <m> } { p i_{μ} }		*

In this tableau, we see that example (55a) is parsed in violation of *$\sigma_{\mu\mu\mu}$, just like the previous schematic example. Candidate (55b) meets *$\sigma_{\mu\mu\mu}$ by leaving a consonant unparsed. Here, *$\sigma_{\mu\mu\mu}$ is the higher ranking of the two constraints, making a tri-moraic hypercharacterized syllable impossible in this situation. See Chapter 3 for a more complete discussion of Finnish.

This situation, where a consonant is left unparsed in the output in preference to parsing it and violating *$\sigma_{\mu\mu\mu}$ seems to be rare. I know of only the Finnish case.

If the three partial rankings in (51) obtain, tri-moraic hypercharacterized syllables closed with a simple moraic consonant are possible in a language. One example of such a language would be Persian (Hayes 1979). Hayes analyzes Persian verse. A great simplification in the description of the versification system is possible if closed syllables with a long vowel are analyzed as tri-moraic.

As part of the factorial typology of this section, we would have a tableau like the following. This is for the example /$a^{\mu\mu}fta^{\mu\mu}b$/, "sun" (Lazard 1992, p. 18).

56) Tableau for Persian Tri-moraic Syllable

		P_{SEG}	*μ_{CONS}	$PARSE_{\mu}$	*$\sigma_{\mu\mu\mu}$
a) ☞	{ $a_{\mu\mu}$ f_{μ} } { t $a_{\mu\mu}$ b }		*		*
b)	{ $a_{\mu\mu}$ <f> } { t $a_{\mu\mu}$ b}	*!			
c)	{ $a_{\mu<\mu>}$ f_{μ} } { t $a_{\mu\mu}$ b}		*	*!	

In this tableau, we see that candidate (56b) is out by $PARSE_{SEGMENT}$, while (56c) is out by $PARSE_{\mu}$. The candidate with the tri-moraic syllable, (56a), is optimal. Note that, in Persian, these syllables are subject to variation. Lazard describes the long vowel in such a syllable as "unstable". This is resolved, when it is resolved, with a $FILL_{\mu}$ violation, not a $PARSE_{SEGMENT}$ violation, as proposed for Finnish in chapter 3, below. Instead, Persian employs a FILL violation, resulting in an epenthetic vowel that takes /f/ as its onset.

There are five rankings that give rise to hypercharacterized syllables closed with a simple moraic consonant. Because we are dealing with four constraints, the number of total rankings is 4! or 24. The other 19 rankings have been discussed in this section as the cases where one of the three partial rankings is not true. Therefore, this is a complete factorial typology of these four constraints.

2.5.1.2 Geminate Hypercharacterized Syllables

Turning now to hypercharacterized syllables closed with a geminate, three constraints are at play, $*\mu_{CONS}$, $PARSE_{\mu}$, and $*\sigma_{\mu\mu\mu}$. The following partial rankings must exist for geminate hypercharacterized syllables to exist.

57a) $PARSE_{\mu}$ >> $*\mu_{CONS}$
b) $PARSE_{\mu}$ >> $*\sigma_{\mu\mu\mu}$

Ranking (57a) is necessary for a language to have geminates in the first place, as discussed earlier in this chapter, after either long or short vowels.

The constraint $PARSE_{\mu}$ must dominate $*\sigma_{\mu\mu\mu}$. Otherwise, there is the possibility of moraic underparsing. Below, I give a schematic tableau of a situation where $PARSE_{\mu}$ dominates $*\sigma_{\mu\mu\mu}$, a situation that has the possibility of geminate hypercharacterized syllables. This is a schematic example, which will be shown with Finnish below.

58) Tableau for $PARSE_{\mu}$ >> $*\sigma_{\mu\mu\mu}$

	$/V^{\mu\mu}G^{\mu}V^{\mu}/$	$PARSE_{\mu}$	$*\sigma_{\mu\mu\mu}$
a) ☞	{ $V_{\mu\mu}$ { G_{μ} } V_{μ} }		*
b)	{ $V_{\mu\mu}$ } { $G_{<\mu>}$ V_{μ} }	*!	

In this tableau, we see that example (58a) is fully parsed in violation of $*\sigma_{\mu\mu\mu}$. Candidate (58b) meets $*\sigma_{\mu\mu\mu}$ by leaving a consonant's mora unparsed. Here, $*\sigma_{\mu\mu\mu}$ is the lower of the two constraints, making a geminate hypercharacterized syllable the more preferred of these two candidates for this language.

If this ranking were reversed, the underparsing of a geminate's mora is predicted. Below, I provide an example from Fula, which is discussed at length in Chapter 4. A similar situation can be found in the language Koya, which is also discussed at length in chapter 4. The resolutions of the two languages are different. As shown below, Fula underparses a geminate's mora. Koya, on the other hand, underparses a long vowel's mora, what we would refer to as closed syllable shortening in a process-oriented system.

59) Tableau for Geminate Mora Underparsing in Fula

	$/Ni^{\mu\mu}w^{\mu}i/$	$*\sigma_{\mu\mu\mu}$	$PARSE_{\mu}$
a) ☞	$\{ N\ i_{\mu\mu} \} \{ w_{<\mu>}\ i_{\mu} \}$		*
b)	$\{ N\ i_{\mu\mu} \{ w_{\mu} \} i_{\mu} \}$	*!	

In this tableau, we see that example (59a) has an unparsed mora and consequently meets $*\sigma_{\mu\mu\mu}$. Candidate (59b) is fully parsed, but consequently fails $*\sigma_{\mu\mu\mu}$. Because $*\sigma_{\mu\mu\mu}$ is the higher of the two constraints, a Geminate may not close a hypercharacterized syllable in Fula (see section 4.2 for the full analysis of Fula).

If the two rankings in (57) are both present in a language, geminate hypercharacterized syllables are predicted. Such a case is Finnish, the focus of Chapter 3. Below I give one example from Finnish, with these three constraints. In this tableau, we see that example (60b) has an unparsed mora and consequently meets $*\sigma_{\mu\mu\mu}$. Candidate (60a) is fully parsed, but consequently fails $*\sigma_{\mu\mu\mu}$. Because $PARSE_{\mu}$ is the higher of the two constraints, geminate hypercharacterized syllables are allowed in Finnish. The ranking of $*\sigma_{\mu\mu\mu}$ and $*\mu_{CONS}$ is not relevant to whether a language may or may not have geminate hypercharacterized syllables. The particular ranking here, however, can be established for Finnish (see section 3.4.3).

60) Tableau for Hypercharacterized Syllable in Finnish

	$/sa^{\mu\mu}k^{\mu}a^{\mu}/$	$PARSE_{\mu}$	$*\sigma_{\mu\mu\mu}$	$*\mu_{CONS}$
a) ☞	$\{ s\ a_{\mu\mu} \{ k_{\mu} \} a_{\mu} \}$		*	*
b)	$\{ s\ a_{\mu\mu} \} \{ k_{<\mu>}\ a_{\mu} \}$	*!		

Geminate hypercharacterized syllables have been analyzed here in terms of three constraints in two partial ranking relationships. There are 3! or 6 possible rankings of these three constraints. Two yield hypercharacterized syllables. The other four have been discussed here as the cases where the rankings are not the same as in (57). Consequently, this is a complete factorial typology of geminate hypercharacterized syllables.

2.5.2 Appendixal Hypercharacterized Syllables

This approach to syllable-final consonants breaks them into appendixes and moraic consonants. A moraic consonant is militated against by $*\mu_{CONS}$. Further, in the event that a moraic consonant follows a bi-moraic vowel, the constraint $*\sigma_{\mu\mu\mu}$ can come into play, as discussed in the previous section. Appendix consonants have a constraint that militates against them, *APPENDIX. As discussed in Chapter 5, there may be separate constraints against syllable final and word-final appendixes, but we will stick to syllable final appendixes here.

There is no special constraint against appendixes after long vowels. If a language has appendix consonants and has long vowels, then it should have hypercharacterized syllables closed with an appendix. Several languages discussed in this dissertation show this. Finnish, which has limited appendixes, has corresponding hypercharacterized syllables closed with appendixes (see Chapter 3). Fula and Koya have appendixes and corresponding hypercharacterized syllables (see Chapter 4). Some Australian languages act this way, including those discussed in 5.2.

As with appendixes in general, what we need to look at are cases where the following partial ranking exists: $PARSE_{SEG}$, $*\mu_{CONS}$ >> *APPENDIX. This ranking insures that a consonant that cannot be parsed as an onset will be parsed as appendixal, not left unparsed or parsed as moraic.

A language with appendixal hypercharacterized syllables is Gumbaynggir, an Australian language. Evidence for the appendixal status of the closing consonant is adduced in 5.2. The further fact that the consonants closing these syllables are of a restricted set is analyzed there and will not be discussed here. The following tableau employs the example /$da^{\mu\mu}lga^{\mu}$/, "sing". What is crucial in this example is the fact that the first syllable has two morae and is followed by a sequence of two consonants at underlying representation.

61) Tableau for Hypercharacterized Syllable in Gumbaynggir

	$/da^{\mu\mu}lga^{\mu}/$	$PARSE_{SEG}$	$*\mu_{CONS}$	*APPENDIX
a)	$\{ d\ a_{\mu\mu}\ l_{\mu} \} \{ g\ a_{\mu} \}$		*!	
b) ☞	$\{ d\ a_{\mu\mu}\ l \} \{ g\ a_{\mu} \}$			*
c)	$\{ d\ a_{\mu\mu}$ <l> $\} \{ g\ a_{\mu} \}$	*!		
d)	$\{ d\ a_{\mu<\mu>}\ l_{\mu} \} \{ g\ a_{\mu} \}$		*!	

Candidate (61a) has a tri-moraic syllable and is out by the constraint $*\mu_{CONS}$. Candidate (61d), which has an unparsed vowel mora, is out by the same constraint. The constraint $*\sigma_{\mu\mu\mu}$ is simply not relevant in avoiding tri-moraic hypercharacterized syllables in this case. Candidate (61c), with an unparsed segment, is out by $PARSE_{SEGMENT}$. The optimal form is (61b), with an appendixal consonant, violating only the low-ranking *APPENDIX. In 5.2, the consonant sequences of Gumbaynggir are analyzed with a sub-hierarchy of *APPENDIX constraints, the lowest of which is the one that appears as *APPENDIX in the tableau above.

2.5.2.1 Stacked Appendix Consonants

This approach to appendixes makes another prediction: that a syllable may have any number of appendix consonants. So long as $PARSE_{SEG}$ is above *APPENDIX, leaving any segment unparsed is worse than parsing it as an appendix. Below, I give a schematic example, with four consonants.

62) Tableau for Stacked Appendix Consonants

	$/V^{\mu}CCCCV^{\mu}/$	$PARSE_{SEG}$	*APPENDIX
a) ☞	{ V_{μ} C C C } { C V_{μ} }		***
b)	{ V_{μ} C C <C> } { C V_{μ}}	*!	**
c)	{ V_{μ} C <C C> } { C V_{μ}}	*!*	*
d)	{ V_{μ} <C C C> } { C V_{μ}}	*!**	

Even though *APPENDIX is violated three times in (62a), that candidate is most harmonic because even one violation of $PARSE_{SEG}$ is fatal when $PARSE_{SEGMENT}$ dominates *APPENDIX.

There are, however, other elements of the grammar which would militate against such sequences. Generalizations about sonority profiles will restrict the number of appendix consonants because sonority can only fall so far. Adjacent consonants are subject to co-occurrence restrictions on place in a variety of languages (Prince 1984, Yip 1991, Lamontagne 1993). Voicing distinctions are restricted in a variety of languages (Lombardi 1991). These and similar restrictions would tend to whittle down such sequences, as would the hiding effects of Broman and Goldstein (1986, 1989).

2.5.2.2 Appendix above $*\mu_{CONS}$

There is one situation where the ranking *APPENDIX >> $*\mu_{CONS}$ will give rise to appendix consonants. This is where the constraint $*\sigma_{\mu\mu\mu}$ is above both *APPENDIX and $*\mu_{CONS}$, when underlying long vowels occur in the input. I have analyzed Finnish along these lines to account for the coronal asymmetry that occurs after long but not short vowels. Below, I show this with a brief tableau for the relevant underlying form. This example is treated at more length in chapter 3, where Finnish is discussed. Here we are concerned only with the fact that the coronal after a long vowel is analyzed as an appendix consonant, not as either a moraic consonant, nor is it left unparsed. This structure is assumed on the basis of coronal asymmetries, which are attributed in this dissertation to an aspect of the appendix constraints (see Chapter 5.)

63) Full Tableau for Finnish Form, /ki$^{\mu\mu}$nte$^{\mu}$/

		*$\sigma_{\mu\mu\mu}$	*APPENDIX$_{COR}$	*μ_{CONS}
a) ☞	{ k i$_{\mu\mu}$ n } { t e$_{\mu}$ }		*	
b)	{ k i$_{\mu\mu}$ n$_{\mu}$ } { t e$_{\mu}$ }	*!		*

In this example, we see that (63a) has a violation of *APPENDIX$_{COR}$, which is above *μ_{CONS}, which is violated by (63b), but (63b) is already out of the running because of the higher ranked *$\sigma_{\mu\mu\mu}$ rules that candidate out. Consequently, the consonant is preferably parsed as an appendix, even with the ranking here.

2.6 Other Underlying Forms

Optimality Theory is concerned with output constraints. As such, there are no constraints on lexical representations (see Prince and Smolensky 1993, especially chapters 5 and 9). Of course, lexical representations may all have certain properties as an effect of the output constraints. The language learner is presented with data that is the output and must create underlying representations based on that data. Consequently, lexical representations may all conform to certain constraints as a result.

In this section I explore two types of inputs which I have not discussed above. First, if a consonant is marked as moraic, we might expect that consonant to occur anywhere in the word. For instance, we might expect a moraic consonant to be word initial, as in the following schematic underlying representation.

64) /G$^{\mu}$V$^{\mu}$.../

On the one hand, there are plenty of languages that disallow initial geminates, even though they have geminates elsewhere in the word. On the other hand, there are languages where there is evidence that there are moraic consonants word initially, at least at underlying representation and sometimes on the surface. Some such consonants surface as geminate when they follow a vowel final prefix. Some languages permit word initial geminates to surface. All of the points made for word initial geminates are also true for word final geminates, in mirror image.

The second type of input we need to consider is the possibility that a moraic consonant might occur preconsonantally. I give a representation below.[13]

65) /...$V^{\mu}C^{\mu}CV^{\mu}$.../

Suppose that this underlying representation exists in a language with the following ranking: $PARSE_{\mu}$ >> $*\mu_{CONS}$ >> $PARSE_{SEG}$. This ranking entails that there will be no consonant sequences, only geminates, as in Woleaian. However, this underlying representation would still be parsed with a consonant sequence.

In section 2.6.1, below, I will discuss the peripheral moraic consonants as in the diagram in (64). I will show there that it is possible to get geminates to simplify at the edges of domains, and it is also possible for geminates to occur at the edges of domains. Section 2.6.2 will be devoted to the preconsonantal moraic consonants. I will propose there that languages with tokens of such underlying representations in their lexicons will be subject to reranking of the relevant constraints.

2.6.1 Peripheral Geminates

An element of the moraic theory of geminates is that marking a simple segment as moraic in the lexicon requires it to be parsed as syllable final, while the requirement that syllables have onsets requires that the segment belong to the following syllable as well. The ambisyllabicity of geminates is thus derived from independent elements of the grammar. Marking a consonant as moraic gives rise to another prediction; that any consonant may be moraic. Consider word initial moraic consonants. Here I take an example from Trukese (Dyen 1965), a Trukese language, related to Woleaian, which has been discussed above. The following is the underlying form of a Trukese word.

66) /$k^{\mu}a^{\mu}n$/ "poi, prepared breadfruit" (p. 7)

When this form occurs in isolation, the initial mora is left unparsed, as in the tableau below. The ranking offered in this tableau is that necessary for geminates, which Trukese has.

---

[13]Inkelas and Cho (1993) make special use of representations like (45); see 2.6.

67) Tableau for PARSE$_\mu$ >> *μ_{CONS} in Trukese

	/$k^\mu a^\mu n$/	PARSE$_\mu$	*μ_{CONS}
a) ☞	{ $k_{<\mu>}$ a_μ n_μ } [14]	*	
b)	{$_{FOOT}$ { k_μ a_μ n_μ } }		*

Candidate (67b) has the mora of the geminate /k/ parsed by the foot, which I assume is a suboptimal parse. I present candidate (67b) as completely shaded on the assumption that it is out by some higher constraint, undominated in Trukese, as discussed below. On this assumption, only (67a) is available for this underlying form.

In Optimality Theory, there are no morpheme structure constraints. However, underlying forms are created by the language learner on the basis of surface structures and the constraint ranking posited by the learner. Because the data available to the learner is constrained by the grammars of the speakers of the previous generation, there may well be underlying forms that will not be posited by the learner. We assume that the language learner will choose the simplest possible surface form. This principle will "constrain" the form of lexical entries, although there are no constraints on underlying representation. Instead, certain underlying forms will never be posited because a simpler underlying representation is available to the learner. This is called Stampean Occultation (Prince and Smolensky 1993, 4.3.1), because the line of thinking seems to have originated with Stampe (1969, 1973).

Consider the situation in (67) from the point of view of Stampean Occultation. Candidate (67a) is preferable, even in this language which has geminates. Further, we would expect no language learner to create an underlying representation like (66) in the first place, all else equal. To see why this is so, consider the following example.

68) Comparison of Underlying Forms and Surface Forms

underlying form	surface form
$k^\mu a^\mu n$	[kan]
$ka^\mu n$	[kan]

Two underlying forms map to identical surface forms. The language learner is then presented with the problem of choosing between the underlying forms. The form which incurs the fewer

[14] The status of the final /n/ is beside the point of our discussion.

violations is the preferred one (see Prince and Smolensky 1993, 4.3.1). The first form will always incur a PARSE$_\mu$ violation, which the second form will never incur, making the second form the preferred underlying form. Therefore, we would expect initial geminates to simplify by Stampean Occultation, all else equal.

What prevents these initial geminates from being reanalyzed as simple segments, as we would expect, is that if they occur after a vowel final word, they are realized as geminates. Below, I give another example, which contains the same morpheme.

69) /je$^\mu$ki$^\mu$si$^\mu$k$^\mu$a$^\mu$n/ "a little poi" (p. 7)

Below is a tableau, with examples parallel to those for the previous, isolation form. I include the irrelevant material in parentheses.

70) Tableau for PARSE$_\mu$ >> *μ_{CONS} with Non-initial Geminate

	/je$^\mu$ki$^\mu$si$^\mu$k$^\mu$a$^\mu$n/	PARSE$_\mu$	*μ_{CONS}
a)	(jeki){ s i$_\mu$ } { k$_{<\mu>}$ a$_\mu$ n}	*!	
b) ☞	(jeki){ s i$_\mu$ {k$_\mu$ } a$_\mu$ n }		*

This case is exactly like the cases of geminates already discussed in section 2.3.2. Candidate (70a) violates the higher ranked PARSE$_\mu$, while meeting *μ_{CONS}. Candidate (70b) violates only the lower *μ_{CONS}. It is preferable to parse the mora, in violation of *μ_{CONS}.

The situation for Trukese, then, is consistent with underlyingly moraic consonants in initial position. In that position, they are parsed only as onsets, not as moraic. After a vowel, they may be parsed as moraic.

In the discussion above, the alternative parse for a mora, as in candidate (67b), was assumed to be less harmonic than parsing a mora by a syllable. This less harmonic parse, however, may be needed to describe the related language, Woleaian. Below, I give an underlying representation of a word form Woleaian. According to Sohn (1975), these initial geminates surface as long. Below, I give a tableau, with a candidate that is a possible structure for a surface, initial geminate, phonetically distinct from simple consonants.

71) /b$^\mu$u$^\mu$lo$^\mu$/

72) Tableau for Woleaian Word-initial Geminate

	$/b^{\mu}u^{\mu}lo^{\mu}/$	$PARSE_{\mu}$	$*\mu_{CONS}$
a)	{ $b_{<\mu>}$ u_{μ} } { l o_{μ} }	*!	
b) ☞	**{**$_{FOOT}$ { b_{μ} u_{μ} } { l o_{μ} } **}**		*

In this tableau, candidate (72b) has a mora parsed by the prosodic word. Morae are preferably parsed by syllables, but this is a possibility that we can consider (see Prince and Smolensky 1993's discussion of the appendix, 4.5). If (72b) is a candidate that GEN can produce, then it will be preferred in this case because it violates only the lower of the two constraints.

Of course, if we admit the possibility of the structure in (72b), we have to ask why Trukese is denied this possible structure. As noted above, the Trukese grammar does not permit this sub-optimal licensing because such candidates are ruled out by a higher constraint, the constraint that morae are optimally parsed by syllables. This constraint is undominated in Trukese, but ranked below $PARSE_{\mu}$ in Woleaian.

2.6.2 Preconsonantal Moraic Consonants

There is no rule or constraint that requires a pre-consonantal consonant to be moraic. The constraint ONSET, which is discussed at length by Prince and Smolensky (1993), requires that syllables have onsets. If a preconsonantal consonant cannot be incorporated into an onset, it must be moraic[15] or it must be left unparsed. If it is left unparsed, the language lacks consonant sequences. If it is parsed as moraic, the language has consonant sequences.

The constraint $PARSE_{\mu}$ is not relevant to this situation. To see why, consider the three candidates below, evaluated according to $PARSE_{\mu}$, $PARSE_{SEG}$, and $*\mu_{CONS}$. Recall that $PARSE_{SEG}$ and $*\mu_{CONS}$ were sufficient for deriving the consonant sequences discussed in section 2.3.1 earlier in this chapter. Furthermore, the derivation of consonant sequences and geminates were completely independent of one another, with neither the existence of consonant sequences implying the existence of geminates, nor the existence of geminates implying the existence of consonant sequences.

[15] We are neglecting appendixes here.

73) Tableau for Example without Moraic Consonant

	$/ta^{\mu}kta^{\mu}/$	$PARSE_{\mu}$	$*\mu_{CONS}$	$PARSE_{SEG}$
a)	{ t a_{μ} k_{μ} } { t a_{μ} }		*!	
b)	{ t a_{μ} $<k>_{<\mu>}$ } { t a_{μ} }	*!		*
c) ☞	{ t a_{μ} <k> } { t a_{μ} }			*

Candidate (73c) is optimal. This is a language which has geminates ($PARSE_{\mu}$ >> $*\mu_{CONS}$), but lacks consonant sequences ($*\mu_{CONS}$ >> $PARSE_{SEG}$), like Woleaian. That candidate (73b) exists is irrelevant. It cannot be the optimal candidate because it is identical to (73b) in all ways, except that it incurs one extra violation of $PARSE_{\mu}$ and will always be less harmonic than (73c).

This brings us to the underlying representation that is the topic of this section, a form like $/ta^{\mu}k^{\mu}ta^{\mu}/$, which has the first of two consonants marked underlyingly as moraic. Below, I give a tableau parallel to the one above. Notice that the three constraints are in the same ranking relationship in the following tableau as they were in the preceding tableau, and the same ranking relationship as posited for the language Woleaian, in section 3.2, above. Recall that Woleaian has geminates, but not consonant sequences, except in very recent loanwords. This ranking relationship was necessary to account for that fact.

74) Tableau for Example with Moraic Consonant

	$/ta^{\mu}k^{\mu}ta^{\mu}/$	$PARSE_{\mu}$	$*\mu_{CONS}$	$PARSE_{SEG}$
a) ☞	{ t a_{μ} k_{μ} } { t a_{μ} }		*	
b)	{ t a_{μ} $<k>_{<\mu>}$ } { t a_{μ} }	*!		*

This tableau differs from the preceding one in that there is no candidate like candidate (73c). Because of the principle of Containment, every candidate from this underlying representation will have the underlying mora from the /k/. Since this is true, the output of this tableau will be (74a). Candidate (74a) has a consonant sequence, even though the ranking of this tableau is $*\mu_{CONS}$ >> $PARSE_{SEG}$, which should have no sequences.

This is a problem for the current approach and, at the same time, an opportunity. It is a problem in the sense that a language with geminates has an underlying representation that will give rise to sequences, without regard to the interaction of $*\mu_{CONS}$ and $PARSE_{SEG}$, contra our earlier claims.

It is an opportunity in the sense that we may want such an underlying representation. Inkelas and Cho (1993) suggest that certain frozen forms of Hausa and Italian are represented as /...C^{μ}C.../, like the forms discussed above. These forms disobey the constraints of these languages on the sonority of the syllable final consonants. In the next subsection, I will sum up Inkelas and Cho's (1993) analysis of Italian, which makes use of prespecification.

2.6.2.1 Prespecification in Italian

The facts of Italian are well-known. Only a sonorant, /s/, or the first half of a geminate may close a syllable. The language Hausa has a similar restriction; I will discuss the better known case of Italian. That a sonorant can close a syllable in Italian can be analyzed, in a process oriented system, as a result of restricting a rule of Weight by Position to apply only to sonorants; Zec (1988, 1992) analyzes Italian this way. The rule of Weight by Position (Hayes 1989, Zec 1988) will assign a mora to a consonant that is followed by another consonant. In this case, only sonorants can receive a mora, following Prince (1983), McCarthy and Prince (1986) and Zec (1988). That /s/ is allowed as well as a sonorant is assumed to follow from the higher sonority of stridents as opposed to other fricatives (see Itô 1986). Following Itô (1989), the fact that a geminate may close a syllable in Italian will follow from the moraic theory of geminates. Since geminate have a mora, there is no need to get one by Weight by Position, an insight that Inkelas and Cho (1993) make greater use of.

Inkelas and Cho (1993) point to a set of counter-examples to the generalization offered. That is, there are some words that break the generalization that only a sonorant may close a syllable. These are comparatively rare and frozen. Such forms, however, must be encoded in the lexicon somehow.

The underlying form for these exceptional words proposed by Inkelas and Cho (1993) are as shown below.

75) Underlying Forms of Italian

μ

kaptare to seek

The heart of this analysis is that these marked forms are permitted to surface with a consonant sequence because the underlying mora is there to license the consonant. If an underlying representation of a consonant sequence were to lack a mora on its first member, that mora could not be provided by the moraification rules of the language.

2.6.2.2 An Optimality Theoretic Analysis of Italian Exceptional Forms

In this section, I will offer an analysis of Italian in Optimality Theoretic terms. This is essentially an adaptation of the approach taken by Inkelas and Cho. The analysis carries easily over to Optimality Theory, although one element does not follow, as noted below. Adopting such an approach allows us to take advantage of the underlying representation with a sequence of the form /...$VC^{\mu}CV$.../.

Consider the following underlying forms for Italian, where I have omitted the vowel's morae for clarity (these examples are from Inkelas and Cho).

76) Underlying Forms in Italian

a)	kaptare (μ over p)	to seek
b)	altro	other
c)	grapa (μ over p)	brandy
d)	cavta	city (nominative, feminine) (surface: catta)

We will need, in translating this analysis to Optimality Theory, to employ a sonority scale for moraification, as in Prince (1983), McCarthy and Prince (1986), Zec (1988, 1992), and Prince and Smolensky (1993, chapter 8). In order to simplify the analysis, I will disregard the /s/ clusters, keeping only to sonorant versus obstruent morae. We will only need a pair of sonority sensitive constraints, as below.

77) Constraints on Sonority of Morae

$*\mu_{OBSTRUENT}$	Obstruents must not be moraic.
$*\mu_{SONORANT}$	Sonorants must not be moraic.

These will be universally ranked with $*\mu_{OBSTRUENT}$ above $*\mu_{SONORANT}$. That is, obstruents universally make worse morae than do sonorants (see Prince and Smolensky 1993, chapter 8, sections 2 and 3.) Introduction of a finer grained sonority scale, along the lines of Itô's (1986) suggestion, will cover the /s/ clusters.

There is one additional aspect to this approach. We will need to adopt Zec's (1993) notion of projection of a mora. That is, for any underlying representation /...VC_iCV.../, there will be a candidate with a mora coindexed with $/C_i/$, , as well as a candidate where the mora, but not the consonant is parsed. I will return to this point below.

We will also need $PARSE_{\mu}$ and $PARSE_{SEGMENT}$.

2.6.2.3 Ranking Arguments

Our first ranking argument involves the constraints $PARSE_{\mu}$ and $*\mu_{OBSTRUENT}$. It can be made on the basis of two forms. First, it can be made with a form with a geminate. This is shown in the ranking tableau below.

78) Ranking Tableau for $PARSE_{\mu}$ and $*\mu_{OBSTRUENT}$

	$/gra^{\mu}p^{\mu}a^{\mu}/$	$PARSE_{\mu}$	$*\mu_{OBSTRUENT}$
a) ☞	{ gr a_{μ} { p_{μ} } a_{μ} }		*
b)	{ gr a_{μ} } { $p_{<\mu>}$ } a_{μ} }	*!	

In this tableau, the first candidate violates the lower of the two constraints, $*\mu_{OBSTRUENT}$, while the second candidate violates the higher of the two constraints, $PARSE_{\mu}$. The two candidates differ in whether or not the geminate's mora is parsed. Since candidate (78a) is the actual form, the ranking relationship must be the correct one.

79a) $PARSE_{\mu}$ >> $*\mu_{OBSTRUENT}$

An identical argument can be made on the basis of the underlying form $/ka^{\mu}p^{\mu}ta^{\mu}re^{\mu}/$, which likewise has a moraic consonant in underlying representation. This is shown in the following ranking tableau, in (80), below. In this tableau, the first candidate violates the constraint $*\mu_{OBSTRUENT}$, while the second candidate violates the higher of the two constraints, $PARSE_{\mu}$. Here the candidates differ in whether the $/p^{\mu}/$ is parsed. Candidate (80b) also incurs a violation of $PARSE_{SEGMENT}$, but this will not invalidate this argument, since it will be established in (81) below that $PARSE_{SEGMENT}$ is ranked below $*\mu_{OBSTRUENT}$. Since candidate (80a) is the actual form, the ranking relationship discussed above must be the correct one.

80) Ranking Tableau for PARSE$_{\mu}$ and *$\mu_{OBSTRUENT}$

	/ka$^{\mu}$p$^{\mu}$ta$^{\mu}$re$^{\mu}$/	PARSE$_{\mu}$	*$\mu_{OBSTRUENT}$
a) ☞	{ k a$_{\mu}$ p$_{\mu}$ } { t a$_{\mu}$ } { r e$_{\mu}$ }		*
b)	{ k a$_{\mu}$ <p$_{\mu}$> } { t a$_{\mu}$ }{ r e$_{\mu}$ }	*!	

Finally, we turn to an underlying form with a sonorant as the first member of a sequence of consonants. Recall that this form is one where a mora that is provided by GEN will be part of the optimal candidate, even though a mora provided by GEN will be less than optimal in the parallel case where a non-sonorant consonant is the first member of the consonant sequence, instead of the sonorant consonant, as we find in the following example.

82) Ranking Tableau for PARSE$_{SEGMENT}$ and *$\mu_{SONORANT}$

	/a$^{\mu}$ltro$^{\mu}$/	PARSE$_{SEGMENT}$	*$\mu_{SONORANT}$
a)	{ a$_{\mu}$ <l>} { tr o$_{\mu}$ }	*!	
b) ☞	{ a$_{\mu}$ l$_{\mu}$ } { tr o$_{\mu}$ }		*

In this tableau, the first candidate, with an unparsed segment, violates the higher of the two constraints, PARSE$_{SEGMENT}$, while the second candidate, which has a moraic consonant, violates the lower of the two constraints, *$\mu_{SONORANT}$.
Since candidate (82b) is the actual form, the ranking relationship must be the correct one.

79b) PARSE$_{SEGMENT}$ >> *$\mu_{SONORANT}$

Another point, which does not constitute a ranking argument, comes from the underlying form /ca$^{\mu}$vta$^{\mu}$/. This part of the analysis is not worked out exhaustively. That is, according to Saltarelli (1970), the result of a concatenation of two morphemes that results in this underlying form maps to a geminate. The two relevant candidates are shown below.

81) Tableau for *$\mu_{OBSTRUENT}$ and PARSE$_{SEGMENT}$

	/ca$^{\mu}$vta$^{\mu}$/	PARSE$_{SEGMENT}$	*$\mu_{OBSTRUENT}$
a)	{ c a_μ v_μ } {t a_μ }		*
b) ☞	{ c a_μ <v> { t_μ } a_μ }	*	*

In this tableau, we see that both candidates violate the constraint *$\mu_{OBSTRUENT}$, while the second candidate violates the constraint PARSE$_{SEGMENT}$. The actual form is (81b), so some higher constraint must rule out (81a). Why a geminate is part of the optimal form is not clear to me. This outcome of Italian phonology is analyzed variously. Inkelas and Cho employ a repair strategy. I see no way to solve it in the current approach at the present time.

These are all the ranking arguments available from these forms of Italian. These will be summed up in the next section.

2.6.2.4 Summary of Ranking Arguments

Below are the rankings established above, with the examples with which they were established and the number of the ranking tableau.

79a) PARSE$_{\mu}$ >> *$\mu_{OBSTRUENT}$ /gra$^{\mu}$p$^{\mu}$a$^{\mu}$/ (78), /ka$^{\mu}$p$^{\mu}$ta$^{\mu}$re$^{\mu}$/ (80)

b) PARSE$_{SEGMENT}$ >> *$\mu_{SONORANT}$ /a$^{\mu}$ltro$^{\mu}$/ (82)

These three rankings are summarized in a diagram below, where a line between two constraints indicates a ranking relationship, with the higher of the two as the higher ranked.

83) Summary of Constraint Rankings

PARSE$_{\mu}$ PARSE$_{SEGMENT}$

*$\mu_{OBSTRUENT}$ *$\mu_{SONORANT}$

The next section provides full tableaux for these forms.

2.6.2.5 Full Tableaux

Below I give full tableaux for the forms discussed above. First is an underlying representation where the first of two candidates has an underlying mora.

84) Full Tableau for /ka$^{\mu}$p$^{\mu}$ta$^{\mu}$re$^{\mu}$/

	/ka$^{\mu}$p$^{\mu}$ta$^{\mu}$re$^{\mu}$/	PARSE$_{\mu}$	PARSE$_{SEG}$	*μ_{OBS}	*μ_{SON}
a) ☞	{ k a$_{\mu}$ p$_{\mu}$ } { t a$_{\mu}$ } {re$_{\mu}$}			*	
b)	{ k a$_{\mu}$ <p$_{\mu}$> } { t a$_{\mu}$ } {re$_{\mu}$}	*!	*		

Next is the form /ca$^{\mu}$vta$^{\mu}$/. This tableau lacks a hand, since the constraints here do not predict the right outcome. Again, the fact that the output has a geminate (instead of a PARSE violation, or, for that matter, a long vowel) is not worked out.

85) Full Tableau for /ca$^{\mu}$vta$^{\mu}$/

	/ca$^{\mu}$vta$^{\mu}$/	PARSE$_{\mu}$	PARSE$_{SEG}$	*μ_{OBS}	*μ_{SON}
a)	{ c a$_{\mu}$ v$_{\mu}$ } { t a$_{\mu}$ }			*!	
b)	{ c a$_{\mu}$ <v$_{\mu}$> } { t a$_{\mu}$ }		*		

Below is a full tableau for the underlying form /gra$^{\mu}$p$^{\mu}$a$^{\mu}$/. The output has a geminate. This is just as we would expect, for a moraic consonant in intervocalic position. Note that this is not the case where the gemination is not explained. That is the previous underlying form.

	/$gra^\mu p^\mu a^\mu$/	PARSE$_\mu$	PARSE$_{SEG}$	*μ_{OBS}	*μ_{SON}
a) ☞	{ gr a_μ { p_μ } a_μ }			*	
b)	{ gr a_μ } { $p_{<\mu>}$ a_μ }	*!			

This tableau is for the underlying form /$a^\mu ltro^\mu$/. The output of this underlying form has a consonant sequence.

87) Full Tableau for /$a^\mu ltro^\mu$/

	/$a^\mu ltro^\mu$/	PARSE$_\mu$	PARSE$_{SEG}$	*μ_{OBS}	*μ_{SON}
a)	{ a_μ <l> } { tr o_μ }		*!		
b) ☞	{ a_μ l_μ } { tr o_μ }				*

As noted above, this analysis does not predict that a form with a geminate is the correct output for the underlying form /$ca^\mu vta^\mu$/. Nonetheless, this analysis has shown how the Inkelas and Cho innovation of marking some non-geminates in the lexicon as moraic can be adapted into an Optimality Theoretic approach, which has been the purpose of this section.

2.6.3 Ranking Occultation

It is possible for a language to be analyzed as having moraic consonants as the first consonant in a sequence underlyingly, as in Italian. The situation described, however, is not a stable one. If a language learner is presented with pre-consonantal moraic consonants, those consonants are prima facia evidence that PARSE$_{SEGMENT}$ is above *μ_{CONS}. Unless the learner has compelling evidence that PARSE$_{SEGMENT}$ is below *μ_{CONS}, learners would simply place PARSE$_{SEGMENT}$ above *μ_{CONS}, thereby analyzing the language as having sequences.

To see why this is reasonable, think of this in terms of the Subset Principle (Pinker 1979, 1989, Wexler and Culicover 1980). Language learners must be able to recognize constraint violations. Further, they must use that evidence to establish the ranking of constraints in their language.[16]

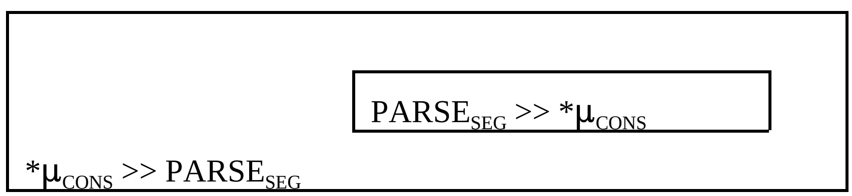
Consider two types of languages, both of which have consonant sequences. The first type, TYPE 1, has PARSE$_{\text{SEGMENT}}$ above $*\mu_{\text{C}}$, just as discussed in 2.3.1. The second type, TYPE 1', has the following ranking of constraints: PARSE$_{\mu}$ >> $*\mu_{\text{C}}$ >> PARSE$_{\text{SEGMENT}}$, as described in 2.6.2. Furthermore, TYPE 1' has underlying representations where the first consonant of an underlying sequence is moraic.

Languages with [CVC] syllables are a subset of languages with only [CV] syllables. That is, all languages which have [CVC] syllables also have [CV] syllables. If we assume the Sub-set Principle, where the default setting of the language acquisition device generates the subset language, then the default ranking will be $*\mu_{\text{CONS}}$ >> PARSE$_{\text{SEGMENT}}$. This is shown as a subset diagram, below.

88) Subset Diagram

no consonant sequences — consonant sequences

$*\mu_{\text{CONS}}$ >> PARSE$_{\text{SEG}}$ — PARSE$_{\text{SEG}}$ >> $*\mu_{\text{CONS}}$



If we assume that children have no access to negative evidence (Brown and Hanlon 1970, Slobin 1972), the ranking they will assume as a starting point is $*\mu_{\text{CONS}}$ >> PARSE$_{\text{SEG}}$. This is the ranking that predicts the subset case, a language which has no consonant sequences. Positive evidence, in the form of consonant sequences, will push language learners to re-rank the constraints as PARSE$_{\text{SEG}}$ >> $*\mu_{\text{CONS}}$. Recall that the two language types, discussed above, TYPE 1 and TYPE 1', both have consonant sequences. Either language as input information to a language learner should result in the re-ranking of the relevant constraints. This is not Stampean Occultation, which was discussed above. Stampean Occultation involves underlying forms being not possible due to the ranking of constraints. I will adopt the term *Ranking Occultation*, since this is a case where a ranking relationship is not possible, due to the learning process.

It is possible that children have access to indirect negative evidence in phonology. That is, since in a very short time, a learner can be exposed to all the possible syllables of a language, the learner might simply decide that any absent syllable type is out. The references in the previous paragraph refer to syntax, where the question of negative evidence is more secure. If it

[16] I say "must" here simply because I can't imagine another way of establishing the non-universal rankings than by studying their interactions.

is the case that children work from this sort of evidence, then the default ranking is not necessary. The reasoning holds up, however. A language with consonant sequences is still evidence for the ranking $PARSE_{SEGMENT} >> *\mu_{CONS}$, with a corresponding simplification of lexical items by omitting the mora on marked forms.

This analysis still leaves the possibility of the Inkelas and Cho (1993) analysis of inadmissible sequences. Learned words and foreign words are presumed not to upset the child's language acquisition process, unless they reach a certain (unknown) threshold, when the result is referred to as "creolization" (see, for instance, Givon (1979), Koefoed (1979).

To sum up, underlying representations of the form $/ta^{\mu}k^{\mu}ta^{\mu}/$ are possible, but subject to reanalysis from one generation to the next. I assume that such forms are not part of the underlying representations of languages, except in a peripheral way.

2.7 Derived Geminates

All of the geminates discussed so far have been marked with a mora in underlying representation. That mora is, due to the principle of Containment, a part of every candidate, although it may be left unparsed in some candidates. We may call these geminates *Underlying Geminates*. A number of languages have consonants that geminate as a result of phonological considerations, where there is no underlying mora. These will be called *Derived Geminates*.[17]

There are two types of Derived Geminates. First, recall the discussion of Woleaian in section 2.3.2, above. That discussion dealt with the fact that Woleaian has geminates. Two candidates from an underlying form without a moraic consonant were considered and are repeated below.

89a) σ_1 σ_2

μ μ

b a s ü

b) σ_1 σ_2

---

[17]Although confusion is nearly inevitable, it should be noted that geminates that are "derived" from affixation of a morpheme with a mora (as can be found in Fula (chapter 4.2.) and in Woleaian (Sohn 1976)) are not derived in the relevant sense. What is of importance here is that these geminates do not associate to an existing mora, rather, to a mora added by GEN. GEN is capable of adding morae, as discussed above.

μ μ μ

b a s ü

In Woleaian, as in other languages discussed here, the first candidate is always less harmonic than the second candidate because the second candidate always includes a violation of $*\mu_{CONS}$. If a language had another constraint that was violated by the first candidate, the second candidate, with its mora provided by GEN, would be more harmonic. Below, the language Wiyot will be discussed.

This will be analyzed as a situation where syllables are optimally bi-moraic, but not with a long vowel.

The second type of derived geminate arises from concatenation. That is, a morpheme ends in a consonant, say /n/, and the next morpheme begins with a token of the same consonant. The result, in the relevant cases, is a geminate. This will be analyzed as a case where Alignment constraints are more highly ranking than OCP constraints.

2.7.1 Derived Geminates in Wiyot

Wiyot (Teeter 1964), a native language of California, does not have underlying geminates.[18] Below is a table with the consonant sequences of Wiyot (p. 13). The table is organized with the first consonant in the column on the left and the second consonant in the row at the top. Capital "L" is a lateral fricative.

[18]This fact will be analyzed below.

Table 1) Consonant Sequences in Wiyot

	t	k	b	s	g	m	n	l	L	y	w
p	✓			✓					✓		
t		✓	✓								✓
k	✓		✓	✓					✓		
k^w	✓								✓		
s		✓	✓								✓
L		✓	✓								✓
h				✓	✓	✓	✓	✓	✓	✓	✓

Wiyot does have surface geminates, as described by Teeter.

"Wiyot syllables always begin with a consonant or cluster followed by a vowel. Where the vowel is short, they must always end phonetically in a consonant, the same as that beginning the next syllable, except at the end of a word. Thus initial and medial syllables are required to be phonetically heavy: . . . where the closing consonantism is not a cluster or /d/, . . . a lengthening of the syllable final consonant is conditioned by this rule."

As noted, Wiyot has word-final light syllables, but no word internal ones, unless the next consonant is /d/, which cannot geminate. The situation here, then, is that syllables must be uniformly bi-moraic, except in two situations where that is not possible.

This distribution of geminates can be analyzed with the following constraints. Some of these constraints are only generally stated below. Some of them are actually parts of families of constraints, which would be formulated in a more insightful and useful manner.

90a) $\sigma_{\mu\mu}$ Syllables must have two morae.

b) *LONG-V Vowels may not be long.

c) *d-gem The segment /d/ may not be geminate.

d) S=L Stressed syllables have long vowels.

e) $*\mu_C$ Consonants must not be moraic.

The first constraint is obviously akin to *$\sigma_{\mu\mu\mu}$. That is, it is a constraint on the moraic content of syllables. It differs from *$\sigma_{\mu\mu\mu}$ in that it is violated when there is only one morae, as well as when there are more than two morae. In Section 2.5.1, the possibility of replacing *$\sigma_{\mu\mu\mu}$ with $\sigma_{\mu\mu}$ was discussed.

The constraint *LONG-V is in opposition to $\sigma_{\mu\mu}$. This constraint is discussed in passing by Prince and Smolensky (1993, section 10.3.4; see also 4.3.3, below.)

The constraint *d-gem states that /d/ is not admissible as a geminate. In analyzing Karok, Bright (1957) posits /d/ as *non-geminable*, just as I am doing here. The reason for its unacceptablility probably follows from the fact that it is a flap. This constraint needs to be part of a theory of the delimitation of geminates and not a specific constraint. This version, however, will serve our purposes.

The constraint S=L governs a relationship between stress and vowel length. Essentially, it is short-hand for the preference for stressed syllables to be more prominent than unstressed syllables. Since vowels are more prominent than are consonants, then the most prominent syllable is one with a long vowel (see Prince (1983), McCarthy and Prince (1986), Hayes (manuscript) and Prince and Smolensky (1993), for discussion of the relationship between prominence and stress; see Prince (1990), with which the current approach is not consistent). This constraint is obviously only one part of the system of constraints that deal with foot well-formedness and the relationship of syllables to feet. Note that the Wiyot examples are dealing with *length*, as opposed to *weight*. Vowel length is not underlyingly distinct in Wiyot. All stress vowels are long. This length "blocks" gemination when an open syllable is stressed (see section 6.3 for some additional discussion of length and weight.)

The constraint *μ_{CONS} is familiar from the discussion of geminates and other moraic consonants in 2.3, above.

2.7.1.1 Ranking Arguments for Wiyot Gemination

Our first ranking argument comes from /sopel/, "in the middle". That argument concerns the constraint $\sigma_{\mu\mu}$ and *μ_{CONS}. In one of the two candidates shown below, there is a mora, provided by GEN, which associates to the onset consonant of the following syllable. The moraic structure of that syllable has been omitted for readability.

91) Ranking Tableau for $\sigma_{\mu\mu}$ and *μ_{CONS}

	sopel	$\sigma_{\mu\mu}$	*μ_{CONS}
a)	{ s o_μ } { p [el]}	*!	
b) ☞	{ s o_μ { p_μ } [el]}		*

In this example, $\sigma_{\mu\mu}$ is violated by the first candidate, which has a mono-moraic syllable. The constraint $*\mu_{CONS}$ is violated by the second candidate, which has a geminate. Because the second candidate is the actual form, this is the correct ranking relationship.

92a) $\sigma_{\mu\mu}$ >> $*\mu_{CONS}$

The next ranking relationship can be made from the same form. It involves the constraints *LONG-V and $*\mu_{CONS}$.

93) Ranking Tableau for *LONG-V and $*\mu_{CONS}$

	sopel	*LONG-V	$*\mu_{CONS}$
a) ☞	{ s o_{μ} { p_{μ} } [el]}		*
b)	{ s $o_{\mu\mu}$ } { p [el]}	*!	

In this example, $*\mu_{CONS}$ is violated by the first candidate. The constraint *LONG-V is violated by the second candidate. The first candidate has a syllable closed with a geminate, while the second has a long vowel. Both meet $*\sigma_{\mu\mu}$. Because the second candidate is the actual form, this is the correct ranking relationship.

92b) *LONG-V >> $*\mu_{CONS}$

Our next ranking argument comes from /sogohw/. It involves the constraints S=L and *LONG-V. In this example, S=L is violated by the first candidate. The constraint *LONG-V is violated by the second candidate.

94) Ranking Tableau for S=L and *LONG-V

	/sogohw/	S=L	*LONG-V
a)	{ s o_{μ} } { g [ohw]}	*!	
b) ☞	{ s $o_{\mu\mu}$ } { g [ohw]}		*

The two candidates differ in the length of the first vowel. Because the second candidate is the actual form, this is the correct ranking relationship.

92c) S=L >> *LONG-V

The next ranking argument can be made on the basis of the form /kadokwsiL/, "he hitches it up". It involves the constraints *d-gem and $\sigma_{\mu\mu}$. Recall that in Wiyot, the segment /d/ will not geminate. The constraint *d-gem militates against a geminate /d/. No doubt, this constraint is actually part of a group of constraints that deal with the segmental content of geminates. In this example, $\sigma_{\mu\mu}$ is violated by the first candidate. The constraint *d-gem is violated by the second candidate. The two candidates differ in whether the /d/ is a geminate. Recall that the constraint *d-gem is not violated in any optimal form in the language. This is following Bright's (1957) insight that the /d/ is *non-geminable*. This non-geminablity is probably related to the fact that /d/ is a flap and that flaps are extremely short, by their very natures, while a geminate is a long segment.

95) Ranking Tableau for *d-gem and $\sigma_{\mu\mu}$

	/kadokwsiL/	*d-gem	$\sigma_{\mu\mu}$
a) ☞	{ k a_{μ} } { d $o_{\mu\mu}$ } [kwsiL]		*
b)	{ k a_{μ} { d_{μ} } $o_{\mu\mu}$ } [kwsiL]	*!	

Since (95a) is the actual form, this is the correct ranking.

92d) *d-gem >> $\sigma_{\mu\mu}$

This form also provides us with our last ranking argument, this time for the constraints *LONG-V and $\sigma_{\mu\mu}$. The constraint *LONG-V militates against long vowels, while the constraint $\sigma_{\mu\mu}$ prefers bi-moraic syllables. In the case of Wiyot, the constraint *LONG-V is the higher ranking of the two, giving us a situation where there will be mono-moraic syllables, when there is no recourse to gemination.

96) Ranking Tableau for *LONG-V and $\sigma_{\mu\mu}$

	/kadokwsiL/	*LONG-V	$\sigma_{\mu\mu}$
a) ☞	{ k a_{μ} } { d $o_{\mu\mu}$ } [kwsiL]		*
b)	{ k a_{μ} } { d $o_{\mu\mu}$ } [kwsiL]	*!	

In this example, the constraint $\sigma_{\mu\mu}$ is violated by the first candidate. The constraint *LONG-V is violated by the second candidate. Just as in the previous case, the difference between the two candidates is the structure of the first syllable. Because the second candidate is the actual form, this is the correct ranking relationship.

92e) *LONG-V >> $\sigma_{\mu\mu}$

These are all the direct ranking arguments that can be made based on these forms. The next section contains a summary of these arguments, as well as transitive rankings.

2.7.1.2 Summary of Rankings

Below is a summary of the rankings discovered above. Each ranking is accompanied with the example from which it can be found and the number from the ranking tableau above.

92a)	$\sigma_{\mu\mu}$	>>	$*\mu_{CONS}$	/sopel/ (91)
b)	*LONG-V	>>	$*\mu_{CONS}$	/sogohw/ (93)
c)	S=L	>>	*LONG-V	/sogohw/ (94)
d)	*d-gem	>>	$\sigma_{\mu\mu}$	/kadokwsiL/ (95)
e)	*LONG-V	>>	$\sigma_{\mu\mu}$	/kadokwsiL/ (96)

These five rankings will not uniquely rank all the five constraints involved in this analysis, either directly or by transitive ranking. These five rankings are summarized in a diagram below, where a line between two constraints indicates a ranking relationship, with the higher of the two as the higher ranked.

97) S=L *d-gem

*LONG-V

$\sigma_{\mu\mu}$

$*\mu_{CONS}$

The only ranking that cannot be established conclusively is the ranking of *d-gem with respect to S=L and *LONG-V. I will put it at the top of the ranking hierarchy, equal to S=L, but not above it, as indicated by separating them with a dashed, not solid, line.

2.7.1.3 Full Tableaux for Wiyot Gemination

Below I include full tableaux for the forms discussed above, as well as the form /sbowi/, from which no distinct ranking argument could be made. The underlying forms, as presented, do not include morae, since vowel length is, in all cases, non-distinctive. The first form, is /sopel/, "in the middle". The optimal output has a geminate, which closes the first syllable. This candidate is preferred over the first candidate, where the first syllable is monomoraic, and the third candidate, where the first syllable contains a long vowel and therefore meets the constraint requiring that all syllables be bimoraic. The first syllable is not stressed, in this example, so the influence of the constraint S=L cannot be felt.

98) Full Tableau for /sopel/, "in the middle"

	sopel	*d-gem	S=L	*LONG-V	$\sigma_{\mu\mu}$	$*\mu_{CONS}$
	{ s o_μ } { p [el]}				*!	
b) ☞	{ s o_μ { p_μ } [el]}					*
c)	{ s $o_{\mu\mu}$ } { p [el]}			*!		

The next tableau shows the underlying form /sogohw/ "it is red".

99) Full Tableau for /sogohw/ "it is red"

	/sogohw/	*d-gem	S=L	*LONG-V	$\sigma_{\mu\mu}$	$*\mu_{CONS}$
a)	{ s o_{μ} } { g [ohw]}		*!		*	
b)	{ s o_{μ} { g_{μ} }[ohw]}		*!			*
c) ☞	{ s $o_{\mu\mu}$ } { g [ohw]}			*		

In the preceding tableau, we see that the optimal form has a long vowel, even in violation of the constraint against bi-moraic vowels.

The next tableau shows the underlying form /kadokwsiL/ "he hitches it up". This form surfaces with a short vowel and no geminate, since those alternatives are both out by higher constraints, in particular, *d-gem and *LONG-V.

100) Full Tableau for /kadokwsiL/, "he hitches it up"

	/kadokwsiL/	*d-gem	S=L	*LONG-V	$\sigma_{\mu\mu}$	$*\mu_{CONS}$
a) ☞	$\{ka_{\mu}\}\{do_{\mu\mu}\}$ [kwsiL]				*	
b)	$\{ka_{\mu}\{d_{\mu}\}o_{\mu\mu}\}$ [kwsiL]	*!				*
c)	$\{ka_{\mu\mu}\}\{do_{\mu\mu}\}$ [kwsiL]			*!		

In the preceding tableau, we see that the optimal candidate has a monomoraic syllable, which is an optimal situation since the alternatives are either a geminate /d/ or a long vowel, both of which are out by higher ranked constraints.

The next tableau is for /sbowi/, "very". In this example, we are concerned with the second of the two syllables. In the following tableau, we see that the optimal candidate is the first one, with a monomoraic vowel in the second syllable. Since that syllable is open and word final, there is no possibility of a candidate with a geminate to close the syllable. It will, consequently, surface as short when it is unstressed.

101) Full Tableau for /sbowi/, "very"

	/sbowi/	*d-gem	S=L	*LONG-V	$\sigma_{\mu\mu}$	$*\mu_{CONS}$
a) ☞	{ sb o_μ { w_μ } i_μ }				*	
b)	{ sb o_μ { w_μ } $i_{\mu\mu}$ }		*!	*		

The only other candidate which meets the requirement that syllables have two morae is in violation of S=L, since that syllable is unstressed, and in violation of the constraint against long vowels.

Similar analyses are possible for the related language of Karok (Bright 1957) and the language Tübatulabal, which has been the focus of much work relating to the question of derived geminates (see Swadesh and Voegelin 1939, McCawley 1969, Lightner 1971, Tranel 1991, Crowhurst 1991).

The fact that Wiyot lacks underlying geminates will follow from Stampean Occultation. Because both simple and moraic consonants map to geminates, the language learner will choose the simpler underlying form. The simpler underlying form of consonants is the form without morae.

This section has included an analysis of derived gemination. This is possible only if there is both a pressure on syllables to be heavy and on vowels to be short. These are embodied in the constraints $\sigma_{\mu\mu}$ and *LONG-V.

2.7.2 Geminates from Concatenation

Another circumstance in which geminates arise is from concatenation of morphemes. That is, a root will end in one token of a consonant and a suffix will begin in another token of the same consonant. English, which lacks underlying geminates, has geminates are the result of concatenation. Some examples involving the suffix *-ness* are listed below.

102) English Geminates from *-ness*

cravenness, crestfallenness, downtroddenness, fineness, forsakenness, goneness, greenness, heartbrokenness, lornness, maudlinness, oneness, saneness, superhumanness, uncommonness

The question of how geminates may occur as the result of concatenation is only part of a larger question. A variety of consonant sequences are found only as a result of concatenation. That is, these sequences are not found within monomorphemic words. For instance, the sequence /tn/ is not found in English monomorphs (Lamontagne 1993). However, it is found over morpheme boundaries. Some examples, again with the suffix *-ness* are given below.

103) Examples with /tn/ from *-ness*

disconsolateness, disparateness, dispassionateness, disproportionateness, explicitness, exquisiteness, extravagantness, fastness, fatness, ferventness, flippantness, justness, lightness, militantness, munificentness, passionateness, peartness, prevalentness, prolateness, proximateness, recentness, requisiteness, rightness, squatness, suppliantness

Our task, then, is to explain why these segments and geminates may occur in polymorphs, but not in monomorphs. This can be done by assuming that intra-morphemic geminates are, in fact, sequences of identical consonants. These geminates, as well as the inadmissible sequences like /tn/, constitute violations of constraints against identical elements being adjacent. The answer offered here is that higher ranking constraints must be met, making the optimal forms have the otherwise inadmissible sequences, including geminates.

In particular, following Lamontagne and Sherer (1993), I am adopting the following alignment constraint for English, first proposed by McCarthy and Prince (1993a) for Axinica Campa.

104) Align (suffix, L, prwd, R)

This constraint states that a suffix is most harmonically attached to a prosodic word. The structure envisioned is a nested prosodic word, as below.

105) Nested Prosodic Word
root $]_{PRWD}$ suffix $]_{PRWD}$ suffix $]_{PRWD}$ .... suffix $]_{PRWD}$

Lamontagne and Sherer (1993) discusses cases where Align is violated in deference to ONSET, but our main concern here is the constraint against identical consonants being adjacent.

106) *CC No two identical consonants may be adjacent.

In fact, this constraint is only the most extreme part of a larger system of dissimilatory effects. No two consonants with the same place of articulation are allowed in English within monomorphemic words (Lamontagne 1993). The constraint *CC, in fact, might simply be *PLACE-PLACE discussed in chapter 5.2.6.

Lamontagne and Sherer (1993) adopt the Alignment constraint in (104). All suffixes, furthermore, subcategorize for a prosodic word. In other words, there is no distinction between stem- and word-selecting suffixes, as in standard Lexical Phonology (Kiparsky 1982, 1985, for selectional distinctions between suffixes, see Selkirk 1982, 1984).

The different behavior of the suffixes follows from the phonological form of the suffixes. Consonant initial suffixes suffix to prosodic word in the optimal candidate. The result is that consonant initial suffixes do not re-arrange the prosodic structure of the word, resulting instead in the nested prosodic word in (105). Therefore, stress, the focus of Lamontagne and Sherer (1993), is not "shifted" by those suffixes.

Vowel initial suffixes, on the other hand, need onsets to satisfy the more highly ranked constraint ONSET. The optimal candidate in such a case will not have the suffix suffixed to prosodic word. Instead, they will be part of a single prosodic word (see McCarthy and Prince 1993a on Axinica Campa). Since the suffixed form has a different prosodic structure form the base, instead of an extra layer of prosodic structure, the stress will fall at a different site than in the unaffixed base.

Here we are concerned with the occurrence of sequences that only occur over morpheme boundaries. In this situation, *CC is subordinated to Align. This ranking gives rise to identical consonants over a prosodic word boundary, as in the following tableau, where *CC is the lowest ranked of the three constraints. It is therefore more harmonic to violate *CC than either ONSET or ALIGN. The symbol] indicates the prosodic word edge, which is referred to in the constraint ALIGN.

107) Tableau for /serene, ness/

		ONSET	Align	*CC
a) ☞	seren]ness			*
b)	seren]<n>ess	*!		
c)	sere<n> ness]		*!	
d)	sere. n <n> ess]		*!	

A language with Align ranked below *CC would be predicted to lack geminates over morpheme boundaries. English has geminates over morpheme boundaries, but lacks them word-internally. The non-occurrence within words is due to the ranking of $*\mu_{CONS}$ and $PARSE_{\mu}$ as discussed in section 3.3.

English lacks geminates that are derived by some affixes, those that are referred to as stem level (Siegal 1972, Aronoff 1976) or Level One in Lexical Phonological models (Kiparsky 1982, 1985). Below, I give some examples.

108) Examples with *in-*, which lack geminates

innate, innoxious, innumerable, innutrition

109) Examples with *un*-, which have geminates

unknit, unknot, unknown, unnatural, unnecessary, unnumbered

The Alignment constraint in (104) is not applicable to (108) and (109), since (104) deals with suffixes. There must be, however, an alignment constraint which deals with prefixes. I offer such a constraint below.

110) Pref Align (prefix, R, prwd, L)

The difference between the two suffixes can now be derived from a representational assumption. If we posit the underlying representation of un- as /un/ and in- as /iN/, where /N/ is a nasal unspecified for place, the difference between the two prefixes will follow from their interactions with a separate constraint, requiring that all segments have a place node, which I call below H-place (for "have place").

To see how this works, consider the underlying form for the surface form "incomplete", /iN, komplit/. This is shown in a tableau below, where [indicates the prosodic word boundary.

111) Tableau for /iN, komplit/

		H-place	Pref
a)	iN [komplit	*!	
b) ☞	[iŋ komplit		*

The prefix /un/ is specified for place. In a parallel tableau, its optimal candidate will not have a violation of H-place, as below, where the example is /un, kommon/.

112) Tableau for /un, kommon/

		H-place	Pref
a) ☞	un [kommon		
b)	[uŋ kommon		*!

That vowel initial words surface with the prefix [in] will follow from coronal markedness (Smolensky 1993).

To return to the examples in (108) and (109), this approach is still inadequate in predicting the geminate versus non-geminate cases. To see why this is true, consider a tableau for the underlying form /iN, numerable/ (consider "numerable" to be morphologically simplex).

113) Tableau for /iN, numerable/

		H-place	Pref
a)	iN [numerable	*!	
b)	[in numerable		*

In the preceding tableau, the better form is (114b), which has two otherwise identical consonants sharing the same place node, which is very reminiscent of the Two Root Theory of geminates, discussed somewhat in chapter 6, below. The constraint *CC should not rule this out, as it is stated above. So, as presented here, this approach will not rule out a geminate in (108).

What this approach can do is offer a structural difference between the (108) cases and the (109) cases. Below is a tableau for the underlying form /un, natural/.

114) Tableau for /un, natural/

		H-place	Pref
a)	un [natural		
b)	[un natural		*!

In this example, (114a) is preferable, with the nested prosodic word structure of (105) in mirror image. If *CC is not allowed to rule out (114a), in deference to Pref, then we would expect a geminate in this form.

The Lamontagne and Sherer (1993) approach offers one solution to the asymmetry between (108) and (109). Another solution to this problem would be to adopt the separate lexical levels into the Optimality Theoretic approach, a tack taken by McCarthy and Prince (1993a) for Axinica Campa. The two lexical levels would be subject to two different rankings of ALIGN and *CC, resulting in one level with and one level without derived geminates.

This other approach not only allows us to adopt wholesale many of the insights of level based models, but could also be used to model variation from register to register in the same grammar. This point has not been the focus of much work in Optimality Theory to date, but see Kiparsky (1993) for some insightful first steps.

In either event, it is possible to predict that a language will have geminates only over morpheme boundaries. This section has been only a brief discussion of such an analysis.

2.8 Conclusions

This chapter has been an exercise in typological predictions. In Optimality Theory, grammars differ primarily, if not exclusively, by the rankings of universal constraints. In this dissertation, I am proposing constraints in certain ranking relationships in order to explain certain phenomena, including the hypercharacterized syllables of Finnish in Chapter 3 and the hypercharacterized syllables of Koya and
Fula in Chapter 4. As such, I am also proposing a typology of languages with respect to certain elements.

The permutations of different constraints discussed here have made several predictions about the distribution of sequences and geminates in the world's languages and the possible hypercharacterized syllables that we may find. To sum up, moraic simple consonants, and consequently the consonants sequences the begin with them, occur when $PARSE_{SEG}$ is above $*\mu_{CONS}$ and fail to arise when $PARSE_{SEG}$ is below $*\mu_{CONS}$. These two constraints are in opposition because parsing a segment as moraic is a violation of $*\mu_{CONS}$. Geminates arise when $PARSE_{\mu}$ is above $*\mu_{CONS}$ and do not arise when $PARSE_{\mu}$ is below $*\mu_{CONS}$. These two constraints are in opposition because parsing a geminate's mora involves violating $*\mu_{CONS}$. Consonant sequences and geminates are not implicationally related to each other because separate PARSE considerations are at play.

Appendix consonants are another source of consonant sequences. Appendix consonants arise only when the constraint against them is ranked below both $*\mu_{CONS}$ and $PARSE_{SEG}$ (but see 2.6.2). When this takes place, the appendix is another possible way of parsing a consonant. Appendix consonants are assumed to be phonetically non-distinct from moraic consonants, so determination of their status must be made by means of prosodic interactions or articulatory asymmetries (see Chapter 5.5).

Hypercharacterized syllables may be closed by either a moraic or an appendixal consonant. In the event they are closed with a moraic consonant, they constitute a violation of $*\sigma_{\mu\mu\mu}$. Only when $*\sigma_{\mu\mu\mu}$ is below certain other constraints can they surface.

When hypercharacterized syllables are closed by appendixal consonants, they are not in violation of $*\sigma_{\mu\mu\mu}$ and consequently, we expect languages with appendixes to have hypercharacterized syllables closed with an appendix. This can happen even in a language with moraic consonants in other environments, as in section 2.4.2.2.

Having considered the broad implications of the approach offered here, we will turn in the next two chapters to individual studies of languages which have some of these patterns.

Sherer, Tim D. 1994 Prosodic Phonotactics. Available from GLSA Publications, Department of Linguistics, South College, University of Massachusetts, Amherst, MA 01003.

PREFACE TO THE COMPUTER-FILE VERSION

This is the third of six chapters of my dissertation. This version of my dissertation is divided into seven chapters, numbered 0 to 6. Chapters 1 through 6 contain the corresponding chapters of my dissertation. Chapter 0 contains the front matter, as well as the appendix and the sources cited.

This version of my dissertation was taken from the Graduate Linguistics Student Association (GLSA) version, which had several typos corrected over the deposit version. This version differs from the GLSA version with respect to some diagrams. Mainly in tableaux, I have substituted bracketing and subscripted morae for tree diagrams in representing syllables, as below:

{C V_μ} {C V_μ}	=	2 cv syllables
{C $V_{\mu\mu}$}	=	a syllable with a long vowel
{C V_μ C_μ}	=	a bi-moraic closed syllable
{C V_μ C}	=	a mono-moraic closed syllable
{C V_μ { C_μ } V_μ }	=	2 syllables with a geminate consonant

There are still some diagrams in their original forms. I hope that these are recoverable from context.

The pagination is different from the original version. Each chapter is paginated separately. There is a table of contents for each chapter at the beginning of that chapter.

This chapter deals with intervocalic consonant sequences of Finnish. A search of a machine readable dictionary revealed a coronal asymmetry in Finnish. Essentially, in Finnish a hypercharacterized syllable may be closed by a geminate or by a coronal consonant, but not by a non-coronal, non-geminate. An Optimality Theoretic analysis of this fact is offered. One aspect of this analysis is that no root not meeting the coronal only requirement would be posited by a Finnish language learner (Stampean Occultation).

Tim D. Sherer
TDSherer@Delphi.Com
December, 1994

FINNISH HYPERCHARACTERIZED SYLLABLES

CHAPTER 3

FINNISH HYPERCHARACTERIZED SYLLABLES

3.1 Introduction

Finnish has two varieties of medial hypercharacterized syllables in its morphemes. Below, are examples of these two types.

1a)	saakka	"until"
b)	kiinte	"real estate"

The first has a long vowel followed by the first half of a geminate. The second has a long vowel followed by the first consonant of a sequence; this consonant must be from a restricted set of consonants, either a coronal or /h/. I analyze the first type as a tri-moraic syllable and the second type as bi-moraic with an appendixal consonant.

In this Optimality Theoretic approach, these two types of hypercharacterized syllables arise as the result of two different constraint interactions. A highly ranked $PARSE_{\mu}$ ensures that a geminate's mora must be parsed, even at the expense of $*\sigma_{\mu\mu\mu}$, in examples like (1a). On the other hand, tri-moraic syllables with moraic simple consonants are less preferred than syllables closed with an appendixal consonant as in (1b).

Aside from the Optimality Theoretic elements of this analysis, an interesting point here is that there is a coronal asymmetry, albeit a limited one. This coronal asymmetry is found only in a certain environment--in a syllable with a long vowel. Coronal asymmetries have been noted in a variety of languages (Fudge 1969, Selkirk 1978 for English, see Paradis and Prunet 1991 for coronal asymmetries from the point of view of underspecification and McCarthy and Taub 1991 and Smolensky 1993). The approach to coronal asymmetries taken here is to attribute them to the appendix. More precisely, there will be several different appendix constraints that refer to the place of articulation of the appendix. These constraint are universally ranked with respect to each other by a phonetic scale (Prince and Smolensky 1993, chapter 5, see also chapter 9).

3.1.1 Tri-moraic Hypercharacterized Syllables

As discussed in chapters 1 and 2, this approach adopts the use of a mora in the lexicon to mark a consonant as a geminate, following McCarthy and Prince (1986, especially 1988) and Hayes (1989) (see Inkelas and Cho (1993) for extensions of this approach that cover inalterability effects). That mora must close a syllable if it is to be parsed and the consonant is to be realized as a geminate. A requirement for onsets, proposed by Itô (1989) and used by Prince and Smolensky (1993, especially chapters 6 and 8), requires association of that consonant as an onset to a following syllable.

The moraic theory of geminates with an Optimality Theoretic implementation is straightforwardly accomplished, as laid out in 2.3.2, above. GEN provides both candidates where the mora is parsed and not parsed, as well as candidates where the melody is and is not an onset. A language which selects as optimal a candidate with a parsed mora, with the melody also parsed as an onset has a geminate as the optimal output for the underlying form of a moraic consonant. The result is a geminate, which is both ambisyllabic and long.

Marking a segment as moraic in the lexicon does raise issues, with respect to hypercharacterized syllables. Recall the discussion in section 2.3.2. Below are schematic examples of possible underlying representations for /...VCV.../ melody sequences, varying with respect to the length of the first vowel and consonant.

2a) VCV	2b) V:CV	2c) VC:V	2d) V:C:V
σ σ	σ σ	σ σ	σ σ
μ μ	μ μ μ	μ μμ	μ μ μ μ
V C V	V C V	V C V	V C V

None of these first three representations violates the bi-moraic limit on the size of a syllable which is captured in the constraint $*\sigma_{\mu\mu\mu}$. In (2d), we have a long vowel followed by a geminate, one of the two types of hypercharacterized syllables in Finnish. While Finnish does allow tri-moraic hypercharacterized syllables, I will show a variety of languages that do not, such Koya and Fula in the next chapter. This is not to say that $*\sigma_{\mu\mu\mu}$ is not real or not a part of the grammar of all languages. The different situations are the result of a conflict between $*\sigma_{\mu\mu\mu}$ and another constraint, PARSE$_{\mu}$, the requirement to include morae as parsed in the output form.

We can see why we would want to keep all the morae from underlying representation, even though we also know that such a mora is not realized in some languages. From a purely functional point of view, the geminate's or a vowel's mora is part of the lexical representation of that segment, and not expressing that length at the surface should carry a cost to the grammar, in the form of difficulties in recoverability. On the other hand, $*\sigma_{\mu\mu\mu}$ is also an element of Universal Grammar and it cannot be violated without cost. We are left with a situation where two elements of grammar come into conflict and one is violated in some language and the other is violated in other languages.

The Optimality Theoretic approach to this variation between languages resides in exploiting constraint conflicts. In essence, this involves having $*\sigma_{\mu\mu\mu}$ and $PARSE_{\mu}$ come into conflict. In languages that allow representations (2d), like Finnish, the requirement that morae be parsed will be higher ranking than $*\sigma_{\mu\mu\mu}$. In languages that do not allow (2d), like Koya and Fula, $*\sigma_{\mu\mu\mu}$ will be the higher ranking constraint.

3.1.2 Bi-moraic Hypercharacterized Syllables

Tri-moraic Hypercharacterized syllables are one of the two types of Hypercharacterized syllables in Finnish. The other type of hypercharacterized syllable in Finnish is a syllable with a long vowel closed with a consonant. The consonant in such an environment is much more restricted than consonants after short vowels. The consonant after a long vowel must be with a coronal or /h/. I am analyzing this effect as following from place restrictions on appendixal consonants. Recall the following diagram from the discussion in Chapter 1.

3) σ_1 σ_2

V_1 C_1 C_2 V_2

In Finnish, C_2 is parsed as the onset of the second syllable, but C_1 cannot be parsed that way. Finnish has no complex onsets, except in loanwords. The consonant C_1 must be licensed as a final element of the preceding syllable, σ_1. There are two options for licensing a consonant in this position, as discussed in chapters 1 and 2. One option is to have the consonant be moraic and the other is to have it be an appendix adjoined to the syllable node with no intervening mora. Finnish exhibits both types of licensing. When V_1, above, is a short vowel, C_1 is moraic. When V_1 is long, C_1 is appendixal (see section 2.4.2.2). This different status between the two cases can account for the coronal asymmetry that Finnish experiences after long, but not short vowels. This analysis will rely on the assumption that appendixes may have place asymmetries, while moraic consonants may not.

3.1.3 Structure of this Chapter

In Section 3.2, I give a brief outline of the source and procedures for sorting the Finnish sequence and geminate data and summarize the main results. A complete description of the data procedure is included, along with all raw counts, in an appendix.

Section 3.3 is a discussion of all the sequence types after long and short vowels. This discussion draws on previous Finnish work to state non-prosodic generalizations about Finnish sequences. It is shown there that the coronal asymmetry does not follow from the other principles governing sequences.

In section 3.4, I discuss the interactions of the constraints. The Finnish data I will analyze is the range of hypercharacterized syllables in the lexicon, with hypercharacterized syllables closed by geminates and by only coronal initial sequences. I will begin by listing the constraints necessary to derive the Finnish data. Then, these constraints will be ranked against each other two at a time. The reader may wish to page ahead to the full tableaux later in the chapter, to have a better feel for the analysis as a whole and to see that the rankings proposed do not trip over each other in the final result. The rankings that we can establish will be summed up and full tableaux given.

Since all underlying forms map to outputs in Optimality Theory, if we want to make statements about the output of a grammar, which is the surface form of the language, then those statements will have to hold of the output of any input whatever. A variety of other inputs are discussed in section 3.5.

Section 3.6 will serve as a conclusion, stating what questions have been answered and what questions are left for other parts of the dissertation.

3.2 The Finnish Data

In this section, I will give the data that are the main focus of this chapter and I will outline my methods for finding the data that is listed here. A fuller explanation is given as an appendix.

Essentially, I created a machine readable lexicon of Finnish words and distilled them down to a list of sequences and geminates, using a variety of computer utilities. A step-by-step description of my procedure is included in the appendix, as well as all the sequence counts, including those that fall below the arbitrary threshold of occurrence.

The list is from Austerlitz's *Finnish Reader and Glossary*, typed by hand by me, with a few minor typos. Austerlitz's morpheme boundaries were included. Finnish orthography was used, except that capitalization was used in place of umlauts (A = ä). Because Finnish orthography is relatively transparent, interpreting these results is much easier and more straight forward than a language like English. The results were spot-checked with a native speaker.

A variety of sources were useful as background on Finnish phonology, particularly Karttunen (1970), Karlsson (1983), Prince (1984), and Keyser and Kiparsky (1984).

The intervocalic geminates and consonant sequences discussed here are those that occur between the first and second vowel of a mono-morphemic form. Monomorphemic forms were used to avoid problems raised by suffixation. Suffixation is governed by a range of alignment constraints which can interfere with the clarity of these generalizations (see McCarthy and Prince 1993a, 1993b, for Alignment constraints in morphology). The data procedure, as described in detail in the appendix, compiled these sequences, along with a diacritic marking whether a sequences followed a long or short vowel. In brief, "murta" became "rt" and "onni" became "nn" on my lists. Examples with long vowels or diphthongs included a diacritic: "taistel" became "=st" and "mietti" became "=tt".

An arbitrary cut-off of four examples was used. That is, if one sequence occurred less than four times, it was considered non-occurring. This move was made to prevent marginal clusters--especially from loanwords--from entering the data. For instance, "ffr" occurred twice in the data, but both were from the same proper name, *affrika*, "Africa."

3.2.1 Results of the Examination

The following tables contain all the sequences of two consonants and all the geminates that occurred four or more times. The first chart has those that occurred after a short vowel and the second chart has those that occurred after a long vowel. Both charts are organized with the first consonant in the column on the left and the second consonant in the row at the top of the table. All sequences are listed in the appendix. For the rest of this chapter, only the *tables* will be referred to.

Table 2) Sequences of Two Consonants after a Short Vowel in Finnish

	p	t	k	s	m	n	l	r	h	j	v	d	g
p	25			4									
t		30	25	8									
k			28	17									
s		50	26	12									
m	21				38								
n		41	17	15		41			9	4		6	15
l	8	16	29		20		42		4	17	28		
r		20		50	22			16	22	32	32		
h		41	13		13			9		9	9	30	

The striking difference between the preceding table and the following table is the narrow range of the sequences occurring after long vowels.

Table 3) Sequences of Two Consonants after a Long Vowel in Finnish

	p	t	k	s	m	n	l	r	h	j	v	d	g
p	8												
t		29		9									
k			43										
s		22	7										
m													
n		13				6							
l		5					9						
r		5											
h		6											

While the lower left portion of the short vowel chart, under the diagonal, is heavily populated, that part of the long vowel chart is sparsely populated and only with sequences containing a coronal or [h]. Geminates are thinner in the second than in the first chart, but include non-coronals. Of course, we would expect fewer tokens after a long vowel than after a short. There were roughly twice as many short vowels in the first syllable in my sample than long vowels. Nonetheless, the differences are profound. Compare short vowel followed by /ks/, 17, with long vowel followed by /ks/, below threshold. There are actually many more /kk/ after a long vowel than after a short.

Section 3.3 is devoted to describing the sequencing possibilities, first after short vowels, then the more specific conditions after long vowels. Geminates will be put aside for now, until we turn to the Optimality Theoretic approach in section 3.4.

3.2.2 Description of Consonant Sequences and Geminates

In this section, I will discuss the patterns found in Finnish sequences. A variety of generalizations dictate the admissible sequences in Finnish. These general conditions apply to sequences after both long and short vowels. In addition to the general conditions on sequences, the sequences found after a long vowel are a subset of those found after a short vowel. The sequences after a long vowel are subject to additional conditions on place and in linking. The conditions on place are a main focus of this chapter. The first member of a sequence must be coronal (or [h], see below). In addition, the sequences are predominately linked structures.

The geminates found after long vowels are not subject to any additional constraints, when compared with the geminates found after short vowels. The only difference between the geminates found in the environment [VV__] and those found in [V__] is that certain derived geminates are not found in the [VV __] environment. These derived geminates, where they are found, are derived from sequences which cannot occur in the VV__ environment. In other words, the geminates which are missing are missing as a result of the lack of certain sequences in the VV__ environment, not as a result of any additional conditions on geminates in general.

This matter of *derived* versus *underlying* geminates is not at odds with the Optimality Theoretic approach. The reason that the consonant sequences are absent from the lexicon is that forms with these sequences are unlearnable in Finnish, as discussed in 3.4.3.2 (see also 2.3.4.1 above, which introduces the notion of Stampean Occultation from Prince and Smolensky 1993 section 4.3.1.) The geminates which do not occur after a long vowel have no source to be derived from, even though they, themselves, are not ruled out in that environment.

While the generalizations that account for the sequences of Finnish are not of interest to us directly, those generalizations require some discussion as a prologue. The contention below is that the difference between sequences after long and short vowels follows from a prosodic effect, specifically the licensing of a consonant in a certain position. In order to precede this way, we must be at least reasonably sure that these differences are not attributable to non-prosodic factors.

The sequences and geminates that occur after a short vowel are listed below. This list is in line with the findings of previous researchers, although my cut-off may cause some differences. I have divided the list and the discussion that follows into categories based on the sonority of the members and separated the geminates from the sequences. Separation of geminates from the rest is a substantive claim, since it is to be expected from the Moraic Theory of Geminates that geminates are independent of consonant sequences.

4a) Sonorant-obstruent Sequences after a Short Vowel

mp nt nk ns nj nd
lp lt lk lj[1] lv
rt rk rs rj rv
ht hk hj hv hd

b) Sonorant-obstruent Sequences after a Long Vowel

nt lt rt ht

5a) Sonorant-sonorant Sequences after a Short Vowel

[1][j] is a glide, but see discussion of these clusters.

lm rm hm lh rh hr nh

b) Sonorant-sonorant Sequences after a Long Vowel
(none)

6a) Obstruent-obstruent Sequences after a Short Vowel

ps tk ts ks st sk

b) Obstruent-obstruent Sequences after a Long Vowel

ts st sk

7a) Geminates after a Short Vowel

pp tt kk ss mm nn ll rr

b) Geminates after a Long Vowel

pp tt kk nn ll

The general sonority scale for consonants will specify liquids as more sonorous than nasals, with obstruents as the least sonorous segments. This is not an unusual move (Selkirk (1978, Steriade (1982)). The scale will be as below.

8) {lr} > {mn} > {-sonorant}

All of the clusters of Finnish fall in sonority, except for the [-son][-son] clusters, to which I will return below.

The two-consonant sequences of Finnish are all inter-syllabic. Except for loans, Finnish does not have complex onsets. Given the sonorants /m n ŋ l r h[2]/ and the obstruents /p t k s/, we would expect the following array of sequences.

9)

mp	nt	ŋk	ns
hp	ht	hk	hs
rp	rt	rk	rs
lp	lt	lk	ls

This prediction is essentially correct. That is, within the general sonority profile specified, the segments combine freely. The one additional feature of this array is that nasal-obstruent

[2]Assuming, here that /h/ is a sonorant. The status of /h/ is a point of contention within phonology and I have nothing to add to that debate.

sequences are homorganic, but this is the usual case in many languages.

Only four sequences from the array are absent from Finnish after short vowels: *[hs], *[hp], *[rp], and *[ls]. I have a general explanation for the first three sequences, following Karttunen (1970), Keyser and Kiparsky (1984), and Prince (1984), but the non-occurrence of [ls] is a mystery to me.

Karttunen notes that [s] and [h] do not sequence together and Keyser and Kiparsky have a rule of debuccalization that affects [s] and yields [h]. This shows, as is common, that [h] and [s] are rather similar segments, differing only in [coronal] (and [sonorant] in the system of Chomsky and Halle 1968). In light of their similarity, the non-occurrence of [hs] can be treated as a dissimilation between two very similar segments.

There are a variety of conditions on /p/ in Finnish, as in Karttunen (1970) (see also Prince 1984), in addition to /p/ being an infrequent sound. I assume that *[hp] and *[rp] are accidental gaps, which are a result of the scarcity of /p/.[3]

The sequences appearing in the list of occurring sequences, but not in the array above are the following: nj, nd, hd, hj, hv, rj, rv, lj, lv. The sequences ending in [j] and [v] are derived from [consonant-k] and [consonant-p], respectively, by a phenomenon known as Gradation in the Finnish literature. Gradation also derives [hd] from [ht]. Therefore, the sequences [hj, hv, rj, rv, lj, lv, hd] will follow from underlying forms that are in the array. The sequences [nj] and [nd] are not the expected outputs of Gradation (see Karttunen for a complete description of Gradation facts). At this time, I have no special account of them, but they fall within my general approach.

After long vowels, fewer sequences are found, both numerically and a narrower range. The second member is always /t/. The first member is always a coronal or /h/. These sequences fall within the pattern followed by those found after short vowels, but are much more constrained. There are also a variety of [+son][+son] sequences, involving r, l, m, n,[4] and h.

The sequences [rm] and [lm] are completely parallel to [rp] and [lp], with both consonants of the sequence licensed by the syllable that it is in. [rn] and [ln] are ruled out by place dissimilation as described in Karttunen.[5] In order to avoid this same dissimilation, [lt] and [rt] must be linked structures.[6] As for the non-occurring [mr] and [ml], I assume that even sequences of [+sonorant][+sonorant] must have a falling sonority profile, as discussed above.

There are no [+sonorant][+sonorant] sequences after long vowels. This is to be expected

---

[3]The lack of clusters involving [p] will not be relevant to the arguments made below. Later in this chapter, coronals and non-coronals will be compared, below, but these comparisons can be made between the velars, alone, and the coronals.

[4]ŋ is derived and occurs only before velars.

[5]Coronal may not be underspecified in Finnish if this generalization is to be captured. Discussing [coronal] as unmarked must be formalized differently, perhaps along the lines of Smolensky 1993.

[6]/lt/ and /rt/ pattern with the [nasal-stop] homorganic clusters with respect to Gradation (see Karttunen), not with the non-homorganic clusters /lp/, /lk/, /rp/, and /rk/.

if the sequences must be linked and if linking is not an available representation for [+sonorant][+sonorant] sequences, as noted above.

Finnish has a variety of [-sonorant][-sonorant] sequences. I am assuming that the lack of [p] in sequences is an accidental gap, as I did for the [+sonorant][-sonorant] sequences above. Assuming that the consonants [p] [t] [k] and [s] are the only obstruents underlyingly in Finnish,[7] the only sequence missing here, once we ignore [p], as above, is [kt].[8] Obstruent-obstruent sequences are not all that common, but there are genuine examples. All contain at least one coronal, as is not uncommon.

The sequences [ts] and [st] apparently disobey the coronal dissimilation that is being discussed above in section 3.3.1.2. I have no special explanation for this, but see Lamontagne (1993) for some discussion of an approach with /st/ as a single complex segment and Padgett (1991) for discussion of dissimilation in Russian, where stops and fricatives from distinct classes, in which dissimilation takes place between members of a class, but not across classes.

The [-sonorant][-sonorant] sequences after a long vowel are restricted to the coronal initial cases. What is not explained by this assumption is the non-occurrence of /tk/ after long vowels, if we assume that the lack of /sp/ is an accidental gap.

Geminates are represented at all points of articulation, after a short vowel. They are also represented at all places after a long vowel. While geminates after long vowels are a subset of geminates after short vowels, this subset relation is not the result of additional constraints on geminates in that environment. The geminates missing after long vowels are [mm], [ss] and [rr]. I assume that the non-existence of [VVrr] [VVss] are accidental gaps[9]. The existence of [ll] and [nn] in the same environment demonstrates that coronal sonorant geminates are possible after a long vowel. The geminate [mm] is derived in Finnish by Consonant Gradation from /mp/. The sequence /mp/ is not admissible in this environment, so the geminate [mm] will have no source from which to be derived.

The segment [h] is sometimes considered to be a placeless consonant, a claim that is also made of coronals (see Paradis and Prunet 1991, and McCarthy and Taub 1991). For instance, Thrainsson (1978) makes crucial use of [h] lacking a consonantal place in his analysis of Icelandic preaspiration. We have evidence that there is a difference between coronals and /h/ in Gradation. Gradation does not treat /ht/ as a linked structure, unlike /nt/ or other [coronal] sequences. Gradation of /ht/ results in /hd/, just as if /t/ were in isolation. The sequence /nt/ and other linked structures result in geminates under Gradation. Therefore, it is not possible to say that /h/ and [coronals] are permitted here for the same reason. A more insightful analysis of the similarities of /h/ and coronals might follow from markedness, as explored for coronals in Smolensky (1993).

---

[7]See Karttunen, who also has an underlying /d/. Whether or not that /d/ is truly underlying will not be relevant here because it does not cluster with obstruents.

[8]Compare this with languages like Greek (Steriade 1982) and English (Lamontagne 1993), where it is usual for the first element of an obstruent-obstruent clusters to be non-coronal.

[9]/rr/ and /ss/ are rarer than other geminates after short vowels in the first place.

To sum up, the sequences which occur after long vowels must meet two requirements. First, their first member is coronal. Second, linking seems to be required. Although this does not lead to a perfect description, considering /ht/ and /sk/, it is clear that neither coronality or linking alone will cover the data. Coronality is required to rule out */VVmpV/ and linking is required to rule out */VVlpV/. Geminates are not subject to the coronality condition, although they are linked structures.

Although I will not discuss this at any greater length than I have done above, it is clear that there are a variety of constraints that restrict the sequences of Finnish after both long and short vowels. I repeat these phenomena here.

10a) an OCP effect on adjacent consonants' articulators

b) a lack of complex onsets

c) a prohibition on rising sonority between consonants

The sequences of Finnish will follow from these principles in some form. When following a long vowel, the sequences are also constrained by prosodic considerations, as in the next section.

3.4 An Optimality Theoretic Account of the Finnish

This section contains an Optimality Theoretic account of the prosodic asymmetries in Finnish. Recall that I will not be discussing the sequence conditions in the preceding section, because they do not differ between the cases when they follow short and long vowels. This section is devoted to the prosodic constraints, which dictate parsing of consonants and morae.

I will proceed as follows. In 3.4.1, I will lay out the forms that need to be derived for Finnish. Section 3.4.2 includes the constraints necessary for Finnish and some discussion of them, especially the nature of PARSE in various constraints. Section 3.4.3 includes ranking arguments, presented in pairwise fashion, for clearest discussion. Section 3.4.4 summarizes the rankings offered. Section 3.4.5 includes full tableaux of the constraints offered.

3.4.1 The Forms Considered for Finnish

We will be concerned with six underlying forms in Finnish. These are the forms with geminates, with non-coronal-initial sequences and with coronal initial sequences. Each of these three types of intervocalic material will be considered after both short and long vowels. Five of these forms will be faithfully parsed. The last form will be parsed unfaithfully so that its underlying form becomes opaic, an instance of Stampean Occultation. The following diagram contains the six underlying forms we will consider. These forms contain the moraic content of the segments.

11) Underlying Representations to be Discussed

	After long vowel	After short vowel
Geminate	/sa$^{\mu\mu}$k$^{\mu}$a$^{\mu}$/ "until"	/i$^{\mu}$k$^{\mu}$u$^{\mu}$na$^{\mu}$/ "window"
Non-coronal initial sequence	/i$^{\mu\mu}$mpi$^{\mu}$/ (contrived)	/ha$^{\mu}$mpa$^{\mu}$/ "tooth"
Coronal initial sequence	/ki$^{\mu\mu}$nte$^{\mu}$/ "real estate"	/a$^{\mu}$nta$^{\mu}$/ "to give"

There are two elements here. First, as mentioned above, there is a cell where a faithful parse of the underlying representation (keeping the length of the vowel and all segments) is not admissible. That form will have to be merged with another surface form.

The other problem involves the status of a coronal after a short vowel. Part of this analysis is that a coronal after a long vowel is appendixal, but we have no reason to think that a coronal is appendixal after a short vowel. In absence of evidence that a coronal is appendixal in that environment, we will assume that it is moraic there and that must be derived in the analysis below. For instance, Kiparsky (1970) makes no note of a difference in the weight of syllables closed with a coronal in his discussion of the meter of the Kalevala.[10]

In Optimality Theory, we are concerned with the output constraints. It should be true that no input should ever have, as its output, one of the non-existent forms of Finnish. Discussing these six forms is not ultimately adequate. These are merely likely suspects. After the main analysis, this point will be addressed.

3.4.2 The Constraints Needed for Finnish

Below is a list of the constraints that are necessary for this Optimality Theoretic account of Finnish syllable weight. This list is given in the ranking which will eventually be argued is correct for Finnish, although the ranking of *APPENDIX$_{NONCOR}$ cannot be fully ascertained. The other ranking relationships will be argued below on the basis of a number of different underlying forms. The ranking arguments will be made with respect to two constraints at a time.

---

[10]Of course, that is not modern Finnish, so we cannot be certain that would be definitive evidence, if it existed.

12a)	*APPENDIX$_{NONCOR}$	An appendix may not be non-coronal.
b)	PARSE$_{\mu}$	Morae must be parsed.
c)	*$\sigma_{\mu\mu\mu}$	Syllables are maximally bi-moraic.
d)	PARSE$_{SEGMENT}$	Segments must be parsed.
e)	*APPENDIX$_{CORONAL}$	Appendixes must not be coronal.
f)	*$\mu_{CONSONANT}$	Consonants must not be moraic.

These constraints have been discussed in chapter 2. However, in Chapter 2, there was a single *APPENDIX constraint, where here there are two. The single *APPENDIX constraint ruled out all appendixal consonants. That discussion is not at odds with the current chapter. The constraint *APPENDIX can simply be viewed as several more specific constraints. They will not appear to be separate unless, as in Finnish, some relevant constraint or constraints separate them. This proposal is discussed in 5.5.

One element of that proposal is a universal ranking of the *APPENDIX constraints by phonetic scale. This means that *APPENDIX$_{NONCOR}$ >> *APPENDIX$_{CORONAL}$ in all languages. This ranking can be established for Finnish independently of that claim, through transitive rankings. That is shown in 3.4.3.1, below.

3.4.3 Ranking Arguments

Below, I give the direct ranking arguments for Finnish. These ranking arguments involve two constraints, two candidates, and a constraint conflict. After each ranking argument and discussion, the two constraints are listed. Transitive ranking, where constraints are ranked by relationship to other constraints, is discussed below.

Our first ranking argument comes from the underlying form /ikkuna/, "window". It involves the constraints PARSE$_{\mu}$ and *$\mu_{CONSONANT}$.

In this tableau, the first constraint must dominate the second constraint, because candidate (13a) violates the second of the two constraints, *$\mu_{CONSONANT}$, while candidate (13b) violates the first of the two constraints, PARSE$_{\mu}$. Notice that (13b) does not violate *$\mu_{CONSONANT}$ because the mora is not associated to parsed segmental material (see 4.2.3). GEN would provide another candidate, of course, which will have that association, violating both constraints and not providing a ranking argument.

13) Ranking Tableau for $PARSE_\mu$ and $*\mu_{CONSONANT}$

	$/i^\mu k^\mu u^\mu na^\mu/$	$PARSE_\mu$	$*\mu_{CONSONANT}$
a) ☞	{ i_μ } { k_μ u_μ } { n a_μ }		*
b)	{ i_μ } { $k_{<\mu>}$ u_μ } { n a_μ }	*!	

Candidate (13a) is the actual outcome, so this ranking relationship must be the correct one.

14a) $PARSE_\mu$ >> $*\mu_{CONSONANT}$

The same ranking argument can be made on the basis of the form $/sa^{\mu\mu}k^\mu a^\mu/$. Below I include a ranking tableau.

15) Ranking Tableau for $PARSE_\mu$ and $*\mu_{CONS}$

		$PARSE_\mu$	$*\mu_{CONS}$
a) ☞	{ s $a_{\mu\mu}$ { k_μ } a_μ }		*
b)	{ s $a_{\mu\mu}$ } { $k_{<\mu>}$ a_μ }	*!	

This ranking would be necessary for any language with geminates, as discussed in Chapter 2.3.2.

The same form, $/sa^{\mu\mu}k^\mu a^\mu/$ "until", which has a long vowel and a geminate. This form has a conflict between $PARSE_\mu$ and $*\sigma_{\mu\mu\mu}$.

16) Ranking tableau for $PARSE_\mu$ and $*\sigma_{\mu\mu\mu}$

	$/sa^{\mu\mu}$ k^μ $a^\mu/$	$PARSE_\mu$	$*\sigma_{\mu\mu\mu}$
a) ☞	{ s $a_{\mu\mu}$ { k_μ } a_μ }		*
b)	{ s $a_{\mu\mu}$ } { $k_{<\mu>}$ a_μ }	*!	

In this tableau, the first constraint must dominate the second constraint, because candidate (16a) violates the second of the two constraints, $*\sigma_{\mu\mu\mu}$, while candidate (16b) violates the first of the two constraints, $PARSE_{\mu}$. Candidate (16a) is the actual outcome, so this ranking relationship must be the correct one.

14b) $PARSE_{\mu}$ >> $*\sigma_{\mu\mu\mu}$

Turning now to the first case involving a sequence, we can use the underlying form /hampaa/ "tooth" to rank the constraints $*APPENDIX_{NONCOR}$ and $*\mu_{CONSONANT}$. In this tableau, $*APPENDIX_{NONCOR}$ must dominate $*\mu_{CONSONANT}$, because candidate (17b) violates the second of the two constraints, $*\mu_{CONSONANT}$, while candidate (17a) violates the first of the two constraints, $APPENDIX_{NONCOR}$. The two candidates differ in whether the /m/ is parsed as an appendixal consonant, as in (17a), or a moraic consonant, as in (17b).

17) Ranking tableau for $*APPENDIX_{NONCOR}$ and $*\mu_{CONSONANT}$

	/ha$^{\mu}$mpa$^{\mu}$/	$*APP_{NONCOR}$	$*\mu_{CONSONANT}$
a)	{ h a_{μ} m } { p a_{μ} }	*!	
b) ☞	{ h a_{μ} m_{μ} } { p a_{μ} }		*

Candidate (17b) is the actual outcome, so this ranking relationship must be the correct one.

14c) $*APPENDIX_{NONCOR}$ >> $*\mu_{CONSONANT}$

Also on the basis of this form, we can rank $PARSE_{SEGMENT}$ and $*\mu_{CONSONANT}$.

18) Ranking tableau for $PARSE_{SEG}$ and $*\mu_{CONSONANT}$

	/ha$^{\mu}$mpa$^{\mu}$/	$PARSE_{SEG}$	$*\mu_{CONSONANT}$
a) ☞	{ h a_{μ} m_{μ} } { p a_{μ} }		*
b)	{ h a_{μ} <m> } { p a_{μ} }	*!	

One way of parsing a consonant in order to meet $PARSE_{SEGMENT}$ is to associate it to a mora, in violation of $*\mu_{CONSONANT}$. We can see that the first constraint must dominate the second

constraint, because candidate (18a) violates the second constraint, $*\mu_{CONSONANT}$, while candidate (18b) violates the first constraint, $PARSE_{SEGMENT}$.
Candidate (18a) is the actual outcome, so this ranking relationship must be the correct one.

14d) $PARSE_{SEGMENT}$ >> $*\mu_{CONSONANT}$

Again, this is what we expect of a language that has moraic consonants. If the ranking were reversed, a consonant would rather be left unparsed, if it could not be parsed as some other structure, than be moraic.

This same ranking can be established on the basis of the underlying form /$a^{\mu}nta^{\mu}$/. This form has a coronal after a short vowel. A coronal a long vowel is optimally parsed as an appendix, as shown in (35) below. We have reason to believe that /anta/ has a moraic /n/. Below, I provide a ranking tableau to illustrate this optimal form.

19) Ranking Tableau for $PARSE_{SEG}$ and $*\mu_{CONS}$

		$PARSE_{SEG}$	$*\mu_{CONS}$
a)☞	{ a_{μ} n_{μ} } { t a_{μ} }		*
b)	{ a_{μ} <n> } { t a_{μ} }	*!	

We will turn now to a non-occurring surface form, "iimpi". We posit an underlying representation, /iimpi/ (contrived). Of course, our task is to find a constraint ranking that never creates iimpi as an output. All other underlying forms should also be mapped to surface forms other than "iimpi". The underlying form /iimpi/ has as its faithful parse "iimpi". This underlying form will be mapped to some surface representation where it merges with a surface representation with a simpler structure. The more complicated underlying representation would not be posited by the language learner. This is an instance of Stampean Occultation. The full course of this will be discussed with the full tableaux below.

This form will permit us to rank the constraints $*APPENDIX_{NONCOR}$ and $PARSE_{SEGMENT}$, as they are presented in this tableau, this time avoiding the non-occurring form in (20a). This point will be returned to in section 3.5, below.)

20) Ranking tableau for *APPENDIX$_{NONCOR}$ and PARSE$_{SEG}$

	/i^{μμ}mpi^{μ}/	*APP$_{NONCOR}$	PARSE$_{SEGMENT}$
a)	{ i$_{μμ}$ m } { p i$_{μ}$ }	*!	
b) ☞	{ i$_{μμ}$ <m> } { p i$_{μ}$ }		*

We see that the first constraint must dominate the second constraint, because candidate (20b) violates the second of the two constraints, PARSE$_{SEGMENT}$, while candidate (20a) violates the first of the two constraints, *APPENDIX$_{NONCOR}$. Candidate (20a) is the less harmonic outcome, so this ranking relationship must be the correct one, with the standard caveat for ranking, especially necessary for this sort of ranking argument, where the actual output is irrelevant, that no other, more highly ranked constraint must be violated by (20a) but not by (20b) and also that there is no third form that is the actual form. Here we argue that a ranking relationship is necessary to avoid an output. This ranking relationship is given below.

14e) *APPENDIX$_{NONCOR}$ >> PARSE$_{SEGMENT}$

A similar argument can be made for the constraints *σ$_{μμμ}$ and PARSE$_{SEGMENT}$.

21) Ranking tableau for *σ$_{μμμ}$ and PARSE$_{SEG}$

	/i^{μμ}mpi^{μ}/	*σ$_{μμμ}$	PARSE$_{SEG}$
a)	{ i$_{μμ}$ m$_{μ}$ } { p i$_{μ}$ }	*!	
b) ☞	{ i$_{μμ}$ <m> } { p i$_{μ}$ }		*

In this tableau, the first constraint must dominate the second constraint, because candidate (21b) violates the second of the two constraints, PARSE$_{SEGMENT}$, while candidate (21a) violates the first of the two constraints, *σ$_{μμμ}$. Candidate (21a) is non-occurring, so the ranking must be *σ$_{μμμ}$ above PARSE$_{SEGMENT}$. The same caveats apply to this argument as held in the last argument.

The reader will have noticed that the first candidate also violates *μ$_{CONSONANT}$ and that violation could, in principle, account for the non-occurrence of (21a). We have already established in (18) that PARSE$_{SEGMENT}$ must dominate *μ$_{CONSONANT}$, so *mora-cons cannot be the reason for (21a)'s unacceptability. Therefore the given ranking relationship must be the correct one.

14f) $*\sigma_{\mu\mu\mu}$ >> $PARSE_{SEGMENT}$

The underlying form /anta/ "to give" establishes a ranking between $*APPENDIX_{CORONAL}$ and $*\mu_{CONSONANT}$.

22) Ranking tableau for $*APPENDIX_{CORONAL}$ and $*\mu_{CONSONANT}$

	$/a^{\mu}nta^{\mu}/$	$*APP_{CORONAL}$	$*\mu_{CONSONANT}$
a)	{ a_{μ} n } { t a_{μ} }	*!	
b) ☞	{ a_{μ} n_{μ} } { t a_{μ} }		*

In this tableau, we see that the first constraint must dominate the second constraint, because candidate (22b) violates the second of the two constraints, $*\mu_{CONSONANT}$, while candidate (22a) violates the first of the two constraints, $*APPENDIX_{CORONAL}$. These two candidates are surface-identical, but we have no reason to believe that [an] is light for the purposes of stress or poetics. Candidate (22b) is the actual outcome, so this ranking relationship must be the correct one.

14g) $*APPENDIX_{CORONAL}$ >> $*\mu_{CONSONANT}$

This underlying form also permits a ranking argument for $PARSE_{SEGMENT}$ and $*\mu_{CONSONANT}$, but that argument was made above.

From the underlying form /kiinte/ "real estate", we can rank $*\sigma_{\mu\mu\mu}$ and $*APPENDIX_{CORONAL}$. This form is analyzed as having an appendixal consonant as part of the coronal asymmetry hypothesis, discussed in chapter 5, below.

23) Ranking tableau for $*\sigma_{\mu\mu\mu}$ and $*APPENDIX_{CORONAL}$

	$/ki^{\mu\mu}nte^{\mu}/$	$*\sigma_{\mu\mu\mu}$	$*APP_{CORONAL}$
a) ☞	{ k $i_{\mu\mu}$ n } { t e_{μ} }		*
b)	{ k $i_{\mu\mu}$ n_{μ} } { t e_{μ} }	*!	

In this tableau, we see that the first constraint must dominate the second constraint, because candidate (23a) violates the second of the two constraints, $*APPENDIX_{CORONAL}$, while candidate (23b) violates the first of the two constraints, $*\sigma_{\mu\mu\mu}$. Candidate (23a) is the actual outcome. As in an earlier case, candidate (23b) also violates $*\mu_{CONSONANT}$, but we have established above the

*APPENDIX$_{CORONAL}$ dominates *$\mu_{CONSONANT}$, so that constraint cannot account for (25b)'s unacceptability. So the given ranking relationship must be the correct one.

14h) *$\sigma_{\mu\mu\mu}$ >> *APPENDIX$_{CORONAL}$

The same form provides us with an argument for the ranking of the constraints PARSE$_{SEGMENT}$ and *APPENDIX$_{CORONAL}$. In this tableau, we see that the first constraint must dominate the second constraint, because candidate (24a) violates the second of the two constraints, *APPENDIX$_{CORONAL}$, while candidate (24b) violates the first of the two constraints, PARSE$_{SEGMENT}$.

24) Ranking tableau for PARSE$_{SEG}$ and *APPENDIX$_{CORONAL}$

	/ki$^{\mu\mu}$nte$^{\mu}$/	PARSE$_{SEG}$	*APP$_{CORONAL}$
a)☞	{ k i$_{\mu\mu}$ n } { t e$_{\mu}$ }		*
b)	{ k i$_{\mu\mu}$ <n> } { t e$_{\mu}$ }	*!	

Candidate (24a) is the actual outcome, so this ranking relationship must be the correct one.

14i) PARSE$_{SEGMENT}$ >> *APPENDIX$_{CORONAL}$

This form also provides us with our last ranking argument, this time between PARSE$_{\mu}$ and *APPENDIX$_{CORONAL}$. In this tableau, we see that the first constraint must dominate the second constraint, because candidate (25a) violates the second of the two constraints, *APPENDIX$_{CORONAL}$, while candidate (25b) violates the first of the two constraints, PARSE$_{\mu}$.

25) Ranking tableau for PARSE$_{\mu}$ and *APPENDIX$_{CORONAL}$

	/ki$^{\mu\mu}$nte$^{\mu}$/	PARSE$_{\mu}$	*APP$_{CORONAL}$
a)☞	{ k i$_{\mu\mu}$ n } { t e$_{\mu}$ }		*
b)	{ k i$_{\mu<\mu>}$ n } { t e$_{\mu}$ }	*!	

Candidate (25a) is the actual outcome, so this ranking relationship must be the correct one.

14j)	PARSE_{μ}	>>	$\text{*APPENDIX}_{\text{CORONAL}}$	

These are all the direct ranking arguments available on the basis of these forms. In the next section these rankings will be summed up and transitive rankings will be discussed.

3.4.3.1 Summary of Constraint Rankings

Below is a summary of the rankings presented above. This list includes the assumed ranking by phonetic scale of the two appendix constraints. With each ranking relationship, I include the form with which it can be established.

14a)	PARSE_{μ}	>>	$*\mu_{\text{CONSONANT}}$	/ikkuna, saakka/ (13), (15)
b)	PARSE_{μ}	>>	$*\sigma_{\mu\mu\mu}$	/saakka/ (16)
c)	$\text{*APP}_{\text{NONCOR}}$	>>	$*\mu_{\text{CONSONANT}}$	/hampa/ (17)
d)	$\text{PARSE}_{\text{SEG}}$	>>	$*\mu_{\text{CONSONANT}}$	/hampa, anta/ (18), (19)
e)	$\text{*APP}_{\text{NONCOR}}$	>>	$\text{PARSE}_{\text{SEGMENT}}$	/iimpi/ (20)
f)	$*\sigma_{\mu\mu\mu}$	>>	$\text{PARSE}_{\text{SEGMENT}}$	/iimpi/ (21)
g)	$\text{*APP}_{\text{CORONAL}}$	>>	$*\mu_{\text{CONSONANT}}$	/anta/ (22)
h)	$*\sigma_{\mu\mu\mu}$	>>	$\text{*APPENDIX}_{\text{CORONAL}}$	/kiinte/ (23)
i)	$\text{PARSE}_{\text{SEG}}$	>>	$\text{*APPENDIX}_{\text{CORONAL}}$	/kiinte/ (24)
j)	PARSE_{μ}	>>	$\text{*APPENDIX}_{\text{CORONAL}}$	/kiinte/ (25)

As noted above, and discussed in Chapter 5, the *APPENDIX constraints are based on a phonetic scale, that of speed of articulation. They are universally ranked according to that scale and the ranking between $\text{*APPENDIX}_{\text{NONCOR}}$ and $\text{*APPENDIX}_{\text{CORONAL}}$ is as below.

26)	$\text{*APPENDIX}_{\text{NONCOR}}$	>>	$\text{*APPENDIX}_{\text{CORONAL}}$

That is to say, non-coronal appendixes are universally worse than coronal ones. In Finnish, $\text{PARSE}_{\text{SEGMENT}}$ can be established to be below $\text{*APPENDIX}_{\text{NONCOR}}$ and above $\text{*APPENDIX}_{\text{CORONAL}}$, providing a transitive ranking argument that the ranking in (26) is correct.

Presented below is a chart which represents the same ranking relationships given above.

On this chart, a line between two constraints indicates an established ranking, with the higher of the two constraints being the higher ranking of the two one. This diagram makes the transitive ranking relationships more obvious than the list above. Here it is clear that the only constraint that cannot be fully ranked is the constraint *APPENDIX$_{NONCOR}$, which can only be ranked above PARSE$_{SEGMENT}$.

27) *APPENDIX$_{NONCOR}$ PARSE$_{\mu}$

*$\sigma_{\mu\mu\mu}$

PARSE$_{SEGMENT}$

*APPENDIX$_{CORONAL}$

*$\mu_{CONSONANT}$

The only rankings that cannot be established by either direct or transitive ranking arguments are the two below.

28a) *APPENDIX$_{NONCOR}$, PARSE$_{\mu}$

b) *APPENDIX$_{NONCOR}$, *$\sigma_{\mu\mu\mu}$

In the full tableaux below, I have set *APPENDIX$_{NONCOR}$ as the highest constraint, equal to PARSE$_{\mu}$, but not ranked with respect to it, as shown by the notation of setting them apart with a dashed line.

3.4.3.2 Full Tableaux

Below are full tableaux for the six underlying representations discussed above. I will provide discussion for the tableaux, which will overlap with the discussion in the preceding section, but by the nature of pairwise argumentation, that discussion was fragmented. In the following tableaux, the constraints are abbreviated for space.

Below is a simple case, where a geminate follows a short vowel. The two constraints in conflict are PARSE$_{\mu}$ and *$\mu_{CONSONANT}$. This case resembles other languages with geminates as discussed in 2.3.2. There isn't much to say about it.

29) Full Tableau--Underlying Form, /$i^{\mu}k^{\mu}u^{\mu}na^{\mu}$/

		$*AP_L$	P_{μ}	$*\sigma_{\mu\mu\mu}$	P_{SEG}	$*AP_C$	$*\mu_C$
a) ☞	{ i_{μ} } { k_{μ} u_{μ} }{ n a_{μ} }						*
b)	{ i_{μ} } { $k_{<\mu>}$ u_{μ} }{n a_{μ} }		*!				
c)	{ i_{μ} } { $k_{<\mu>}$ u_{μ} }{n a_{μ} }		*!				

Because there are only two morae in this syllable in any of these candidates, the constraint against tri-moraic syllables will not come into play.

Candidates (29a) and (29b) differ in whether the geminate's mora is parsed. Parsing the geminate's mora in (29a) violates $*\mu_{CONS}$, while the unparsed mora in (29b) is a violation of PARSE$_{\mu}$. With the ranking offered, the candidate with a parsed mora is more harmonic.

Another conceivable candidate is shown below.

30) Alternative Candidate

σ σ σ

μ μ μ μ

i k u n a

In this candidate, the geminate's mora is parsed to the vowel melody. This is what would be called compensatory lengthening in a process oriented system. In Finnish, the constraint Segmental Integrity must be undominated. This rules out candidates with elements associating the segments other than those of which it is a part underlyingly.

I have discuss the candidate in (29c) in passing before. Under that recursive definition of PARSE, the segment is not parsed as a moraic consonant, thus there is no extra violation of $*\mu_{CONS}$.

Below we have a case where the first syllable has a long vowel and a geminate.

31) Full Tableau--Underlying Form, /$sa^{\mu\mu}k^{\mu}a^{\mu}$/

		$*AP_L$	P_μ	$*\sigma_{\mu\mu\mu}$	P_{SEG}	$*AP_C$	$*\mu_C$
a) ☞	$\{s\ a_{\mu\mu}\ \{\ k_\mu\ \}a_\mu\ \}$			*			*
b)	$\{s\ a_{\mu\mu}\ \}\{k_{<\mu>}\ a_\mu\}$		*!				
c)	$\{s\ a_{\mu<\mu>}\{k_\mu\ \}a_\mu\ \}$		*!				*

Because the long vowel has two morae and the geminate has one, the faithfully parsed syllable in (31a) has three morae, in violation of $*\sigma_{\mu\mu\mu}$. This violation is preferable to leaving unparsed either a vowel mora, as in (31c) or a consonant mora, as in (31d), although parsing the mora gives rise to another violation, as can be seen from this tableau. This effect is discussed in section 4.2.2.

Below is a form with a short vowel and a consonant sequence. The consonant sequence does not begin with a coronal. In a following case, there will be an example where the consonant sequence does begin with a coronal.

Here we find that a moraic consonant is the optimal parse, in (32b). The candidates (32a) and (32c) contain violations of higher constraints, $*APPENDIX_{NONCOR}$ and $PARSE_{SEGMENT}$.

32) Full Tableau--Underlying Form, /$ha^{\mu}mpa^{\mu}$/

		$*AP_L$	P_μ	$*\sigma_{\mu\mu\mu}$	P_{SEG}	$*AP_C$	$*\mu_C$
a)	$\{h\ a_\mu\ m\}\ \{p\ a_\mu\ \}$	*!					
b) ☞	$\{h\ a_\mu\ m_\mu\}\ \{p\ a_\mu\}$						*
c)	$\{h\ a_\mu\ <m>\}\ \{pa_\mu\}$				*!		

This tableau provides us with the ranking arguments for $*APPENDIX_{NONCOR}$ and $PARSE_{SEGMENT}$ being above $*\mu_{CONSONANT}$. Any language with moraic non-geminates will have to have this ranking, as discussed in Chapter 2. If either an appendix constraint or $PARSE_{SEGMENT}$ is below $*\mu_{CONSONANT}$, then the first segment of a two consonant sequence will be an appendix or it will remain unparsed.

Below is a case with a long vowel followed by a consonant sequence at underlying representation. This consonant sequence begins with a non-coronal.

This case involves mapping the underlying form /iimpi/ to an unfaithful surface form, in this case [iipi]. We know that this is the correct form because of an independently established ranking between $PARSE_{\mu}$ and $*\sigma_{\mu\mu\mu}$, which would not allow us to avoid the illicit surface form of candidate (33b), if $PARSE_{SEGMENT}$ where above $PARSE_{\mu}$.

33) Full Tableau--Underlying Form, $/i^{\mu\mu}mpi^{\mu}/$

		$*AP_L$	P_{μ}	$*\sigma_{\mu\mu\mu}$	P_{SEG}	$*AP_C$	$*\mu_C$
a)	$\{ i_{\mu\mu}\ m \} \{ p\ i_{\mu} \}$	*!					
b)	$\{ i_{\mu\mu}\ m_{\mu} \} \{ p\ i_{\mu} \}$			*!			*
c) ☞	$\{ i_{\mu\mu} <m> \} \{ p\ i_{\mu} \}$				*		
d)	$\{ i_{\mu<\mu>}\ m \} \{ p\ i_{\mu} \}$		*!				*

Both (33b) and (33d) violate $*\mu_{CONSONANT}$, but it has been established clearly that the constraint $*\mu_{CONSONANT}$ must be below $PARSE_{SEGMENT}$, so those violations cannot be what rules out candidates (33b) and (33d).

The reason that Finnish lacks these forms, then, is that such forms are not learnable, by Stampean Occultation (Prince and Smolensky 1993, 4.3.1 and also chapter 9; this line of thinking is due to Stampe (1969, 1973). That is, because the underlying $/i^{\mu\mu}mpi^{\mu}/$ is always mapped to surface [iipi], the language learner will not posit $/i^{\mu\mu}mpi^{\mu}/$ as an underlying form because the simpler $/i^{\mu\mu}pi^{\mu}/$ is available. Consequently, there are no roots of the form $/i^{\mu\mu}mpi^{\mu}/$.

Below is a short vowel followed by a coronal initial consonant sequence.

34) Full Tableau--Underlying Form, /a$^{\mu}$nta$^{\mu}$/

		*AP$_{L}$	P$_{\mu}$	*$\sigma_{\mu\mu\mu}$	P$_{SEG}$	*AP$_{C}$	*μ_{C}
a)	{ a$_{\mu}$ n } { t a$_{\mu}$}					*!	
b) ☞	{ a$_{\mu}$ n$_{\mu}$ } { t a$_{\mu}$}						*
c)	{ a$_{\mu}$ <n> } { t a$_{\mu}$}				*!		

The main point of interest here is the fact that *APPENDIX$_{CORONAL}$ is ranked above *$\mu_{CONSONANT}$. This is to account for the fact that there is no reason to believe that a coronal is not moraic after a short vowel. Indeed, there is no mention in the poetic system of Finnish (for instance, Kiparsky 1970) that syllables with a short vowel followed by a coronal sequence are light. This situation can be captured with the ranking given.

In the following tableau, however, we have the case where a long vowel is followed by a coronal. In this case, this analysis does posit a non-moraic consonant for the closing consonant in the hypercharacterized syllable. Note that this is a representational solution to the problem of why the syllables are different. That is, the appendix has a different structure than do moraic consonants and are therefore not subject to the same constraint against tri-moraic syllables. The coronal asymmetries are posited to follow from a fact about appendixes, that appendix constraints differentiate between different places of articulation (see 5.5, below).

35) Full Tableau--Underlying Form, /ki$^{\mu\mu}$nte$^{\mu}$/

		*AP$_{L}$	P$_{\mu}$	*$\sigma_{\mu\mu\mu}$	P$_{SEG}$	*AP$_{C}$	*μ_{C}
a) ☞	{ k i$_{\mu\mu}$ n}{ t e$_{\mu}$}					*	
b)	{ k i$_{\mu\mu}$ n$_{\mu}$}{ t e$_{\mu}$}			*!			*
c)	{ k i$_{\mu\mu}$ <n>}{ t e$_{\mu}$}				*!		
d)	{ k i$_{\mu<\mu>}$ n$_{\mu}$}{ t e$_{\mu}$}		*!				*
e)	{ k i$_{\mu<\mu>}$ n}{ t e$_{\mu}$}		*!			*	

Above is the case where a coronal initial consonant sequence closes a hypercharacterized syllable. Notice that this cannot be a case where the two mora limit on syllable size is violated, because we would also expect inadmissible [iimpi] forms, because there are no constraints on place imposed by morae. This form shows us how the phonetic scale approach to the appendix constraints works. When other constraints, this time $\text{PARSE}_{\text{SEGMENT}}$ and $*\sigma_{\mu\mu\mu}$ intervene between the two appendix constraints, two different behaviors will arise between coronals and non-coronals, just as happens here.

A non-coronal appendix is less preferred than an unparsed segment. At the same time, an unparsed segment is less preferred than a coronal appendix. If the two appendix constraints were adjacent in the hierarchy, or if no relevant constraint separated them, there would be no coronal asymmetry. This is the case in Koya, in the next chapter. That the asymmetry will always favor parsing coronals follows from the universal ranking of $*\text{APPENDIX}_{\text{NONCOR}}$ above $*\text{APPENDIX}_{\text{CORONAL}}$. This case also has the attribute of parsing a consonant as appendixal in one environment when it is parsed as moraic in another environment in the same language. This point was discussed in 2.5.2.2, above.

3.5 Other Underlying Forms

A goal of Optimality Theory is to create a grammar that maps all possible underlying forms to grammatical surface forms of the language (see Prince and Smolensky 1993, especially chapters 1 and 2). In Finnish, above, a number of candidates were faithfully parsed to admissible forms of the language and one underlying form, /$i^{\mu\mu}mpi^{\mu}$/, was unfaithfully parsed to an admissible form. A faithful parse, one that maps a form to a surface form without unparsed material or empty structure, is not the only possible source of an inadmissible form. If we have our analysis wrong, it could predict that some other form would be unfaithfully parsed to that inadmissible form. Here I will discuss some other underlying forms that could be mapped to the inadmissible surface [iimpi]. The goal here is to show that some other parse of these forms, often but not necessarily a faithful parse, is preferable to the inadmissible [iimpi].

Our goal is to find out if some underlying representation creates a hypercharacterized syllable closed with a consonant sequence which does not begin with a coronal. One possible representation is offered below, with the relevant candidate to the left.

36) /$i^{\mu\mu}mi^{\mu}pi^{\mu}$/ iim<i> pi

This case involves a PARSE violation. A vowel is left unparsed, leaving /m/ to be associated, either as moraic or appendixal, to the first syllable. For such a mapping to take place, the violation of $\text{PARSE}_{\text{SEGMENT}}$ must not prove fatal to this candidate. In the following tableau, we see that it does.

The first candidate is a faithful parse and violates none of these constraints. The second candidate has an unparsed vowel and a moraic /m/. In this candidate, the /m/'s mora comes from that underlying vowel in violation of Segmental Integrity, but we needn't include that to see why this form is not preferred. It incurs three violations of constraints, while candidate (36a) incurs none. Candidate (36c) has the /m/ parsed as appendixal. This incurs the violations of three constraints, making it less harmonic than the first candidate.

37) Full Tableau--Underlying Form, /$i^{\mu\mu}mi^{\mu}pi^{\mu}$/

		*AP_L	P_μ	*$\sigma_{\mu\mu\mu}$	P_{SEG}	*AP_C	*μ_C
a)	{$i_{\mu\mu}$ }{mi_μ}{p i_μ}						
b)	{$i_{\mu\mu}$ m<i_μ>}{p i_μ}			*!	*		*
c) ☞	{$i_{\mu\mu}$m<i_μ>}{p i_μ}	*!	*		*		

Therefore, the underlying form /$i^{\mu\mu}mi^{\mu}pi^{\mu}$/ will map to something other than [iimpi]. Whether that is actually candidate (36a) is not relevant, since we know that at least candidate (36a) is better than (36b) or (36c).

We might also consider a case where there were three consonants in a sequence in underlying representation, such as /$i^{\mu\mu}mtpi^{\mu}$/. This is obviously out because it will always be more harmonic to leave extra consonants unparsed than to parse them as either moraic or appendixal. To parse a segment as moraic in such a case involves a violation of the more highly ranked *$\sigma_{\mu\mu\mu}$ and to parse a non-coronal as appendixal involves a violation of the more highly ranked *$APPENDIX_{LAB}$. The same will be true no matter how many consonants are in the sequence underlyingly, because absolute domination will require any number of $PARSE_{SEGMENT}$ violations in order to meet the more highly ranked *$\sigma_{\mu\mu\mu}$ or *$APPENDIX_{LAB}$.

Another possible underlying form that could map to the inadmissible [iimpi] would be /$i^{\mu\mu}pi^{\mu}$/. This mapping would involve a violation of $FILL_\mu$. This constraint has been discussed in passing in 2.3 and 2.3.1. I repeat it below.

38) $FILL_\mu$ Morae must be associated to segmental material.

A violation of this constraint could give rise to an optimal candidate [iimpi]. This is shown in the tableau below for an underlying form /$i^{\mu\mu}pi^{\mu}$/. As can be seen from this short tableau, candidate (38b), which is the form with a hypercharacterized syllables, violates both *$\sigma_{\mu\mu\mu}$ and *μ_{CONS}, where candidate (38a) violates neither.

We needn't concern ourselves with where in the hierarchy $FILL_\mu$ is ranked. Whether it is above or below either or both of these constraints, it cannot save (38b), since (38b) violates $FILL_\mu$. Whether that violation is above or below the ones in the following tableau is not relevant.

39) Tableau--Underlying Form, $/i^{\mu\mu}pi^{\mu}/$

		$*\sigma_{\mu\mu\mu}$	$*\mu_{CONS}$
a)	$\{ i_{\mu\mu} \} \{ p\, i_{\mu} \}$		
b)	$\{ i_{\mu\mu}\, \square_{\mu} \} \{ p\, i_{\mu} \}$	*!	*

We might ask if the epenthetic segment could be a coronal, thereby avoiding the violations in (38) above, in favor of a violation of $*APPENDIX_{CORONAL}$. This move would not create an inadmissible syllable, as such syllables are possible. Nevertheless, this would not be the optimal candidate, since it incurs violations of both $*APPENDIX_{CORONAL}$ and $FILL_{\mu}$, while (38a) incurs neither violation.

So far in this section, we have considered candidates where underparsings or overparsings create candidates that are not admissible forms of Finnish. In each case, at least one more harmonic candidate can be found. This means that the non-occurring surface forms will not occur because they are not optimal. There is one underlying form which could map to an inadmissible form which does not involve an unfaithful parse: $/i^{\mu\mu}m^{\mu}pi^{\mu}/$.

40) Tableau--Underlying Form, $/i^{\mu\mu}m^{\mu}pi^{\mu}/$

		$*AP_{L}$	P_{μ}	$*\sigma_{\mu\mu\mu}$	P_{SEG}	$*\mu_{C}$
a) ☞	$\{ i_{\mu\mu}\, m_{\mu} \}\{ p\, i_{\mu} \}$			*		*
b)	$\{ i_{\mu\mu} <m_{\mu}> \}\{ p\, i_{\mu} \}$		*!		*	
c)	$\{ i_{\mu\mu}\, m_{<\mu>} \}\{ p\, i_{\mu} \}$	*!	*			
d)	$\{ i_{\mu\mu\mu} <m> \}\{ p\, i_{\mu} \}$			*!	*	*

This candidate has the first consonant of an underlying sequence marked as moraic underlyingly. Candidates where the first member of a sequence are marked as moraic underlying were discussed in 2.6.2, following Inkelas and Cho (1993). It was concluded there that such representations were not only possible sources of consonant sequences, but sometimes desirable ones. We might ask if this general result also holds in the case of Finnish. Below is a tableau for this underlying form. I have omitted one of the appendix constraints for space. As with the cases discussed in 2.6.2, I will analyze this as a case of Ranking Occultation. That is, if there were available surface forms like [iimpi], the language learner would rerank the constraints so that

$PARSE_{SEG}$ dominates $*\sigma_{\mu\mu\mu}$ in order to simplify underlying representation.

Underlying representations on a par with $/i^{\mu\mu}m^{\mu}pi^{\mu}/$ are, of course, possible for the few forms that exceptionally have this inadmissible surface form, such as juokse, "current", tuoksu, "good odor", vuoksi, "because of, on account of", and kuinka, "how" (glosses from Austerlitz 1965).

3.6 Summary

Finnish provides us with several points of interest, but none so fundamental to our understanding of hypercharacterized syllables as the interaction of the maximal syllable size constraint, $*\sigma_{\mu\mu\mu}$, with other constraints. The constraint $*\sigma_{\mu\mu\mu}$ conflicts crucially with $PARSE_{\mu}$, $PARSE_{SEGMENT}$, and $*APPENDIX_{CORONAL}$. When $PARSE_{\mu}$ dominates $*\sigma_{\mu\mu\mu}$, as it does in Finnish, the language has the possibility of hypercharacterized syllables which are tri-moraic and will have the form $[...VVG]_{\sigma}$, where G is the first half of a geminate (providing, of course, that $PARSE_{\mu}$ also dominates $*\mu_{CONSONANT}$, so that the language may have geminates in the first place). This analysis is desirable because it makes a natural cut between geminates and non-geminate consonants, based on the difference between them--a geminate has a mora underlyingly and non-geminates do not, as a result of the moraic theory of geminates. The analysis does not have to "look back" at the lexicon in order to do this, as we have seen above.

The question of what happens when $*\sigma_{\mu\mu\mu}$ dominates $PARSE_{\mu}$ is taken up in Chapter 4, where the two languages Koya and Fula present cases where parsing a mora is less important than maintaining the syllable size template.

Also of interest in the Finnish case is the interaction of $*\sigma_{\mu\mu\mu}$ with the appendix constraints. The Appendix constraints will not treat different consonants differently unless the constraints are separated by some other constraint. In the Finnish case above, segments were kept from being licensed as appendixes by the constraint $*APPENDIX_{NONCOR}$. Since that constraint is above $*\sigma_{\mu\mu\mu}$ and other constraints, some other candidate would be optimal. If a candidate violates $*APPENDIX_{CORONAL}$, on the other hand, $PARSE_{SEGMENT}$ and $*\sigma_{\mu\mu\mu}$ are higher and rule out the other candidates, leaving the candidate with an appendix preferred. Appendix consonants, including their ability to distinguish between different places of articulation and how this follows from constraints separating the appendix constraints, are the main point of Chapter 5, but the existence of appendix consonants, as defined here, is argued for in Chapter 4.

Sherer, Tim D. 1994 Prosodic Phonotactics. Available from GLSA Publications, Department of Linguistics, South College, University of Massachusetts, Amherst, MA 01003.

PREFACE TO THE COMPUTER-FILE VERSION

This is the fourth of six chapters of my dissertation. This version of my dissertation is divided into seven chapters, numbered 0 to 6. Chapters 1 through 6 contain the corresponding chapters of my dissertation. Chapter 0 contains the front matter, as well as the appendix and the sources cited.

This version of my dissertation was taken from the Graduate Linguistics Student Association (GLSA) version, which had several typos corrected over the deposit version. This version differs from the GLSA version with respect to some diagrams. Mainly in tableaux, I have substituted bracketing and subscripted morae for tree diagrams in representing syllables, as below:

{C V_{μ}} {C V_{μ}}	=	2 cv syllables
{C $V_{\mu\mu}$}	=	a syllable with a long vowel
{C V_{μ} C_{μ}}	=	a bi-moraic closed syllable
{C V_{μ} C}	=	a mono-moraic closed syllable
{C V_{μ} { C_{μ} } V_{μ} }	=	2 syllables with a geminate consonant

There are still some diagrams in their original forms. I hope that these are recoverable from context.

The pagination is different from the original version. Each chapter is paginated separately. There is a table of contents for each chapter at the beginning of that chapter.

This chapter includes analyses of two languages, Fula and Koya, both of which permit long vowels before sequences of consonants, but not before geminates. This is analyzed by means of a syllable appendix--a consonant directly adjoined to the syllable node without an intervening mora. This version of an appendix consonant differs from the word-final version. Both Fula and Koya have cases where we might expect long vowels to precede geminates. In Fula, these cases surface with a simple consonant (corresponding to the geminate), while in Koya, these forms surface with a short vowel. Optimality Theoretic analyses for these mappings are given.

Tim D. Sherer
TDSherer@Delphi.Com
December, 1994

HYPERCHARACTERIZATION IN FULA AND KOYA

CHAPTER 4

HYPERCHARACTERIZATION IN FULA AND KOYA

4.1 Introduction

In the previous chapter, we saw that Finnish permits hypercharacterized syllables with a long vowel followed by a geminate. As discussed in Chapter 2, this situation arises only when the relevant constraints, in this case PARSE$_{\mu}$, *μ_{CONS}, and *$\sigma_{\mu\mu\mu}$, are in a certain ranking relationship. Essentially, PARSE$_{\mu}$ must dominate both *μ_{CONS} and *$\sigma_{\mu\mu\mu}$. It's good that we can get this result, but it isn't enough. In order to have a theory of the distribution of hypercharacterized syllables across languages, we need not only to be able to describe how a language may have them, but also under what circumstances a language *may not* have hypercharacterized syllables. Previous work on hypercharacterized syllables has not met this need. The usual answer to why one language permits and another does not permit hypercharacterized syllables is along the lines of a parameter; there is a principle, such as a constraint or a persistent rule, which prohibits hypercharacterized syllables. In some languages, it is off, in others it is on. Even this answer runs into problems. If a language has hypercharacterized syllables closed with consonant sequence and has geminates, shouldn't it have hypercharacterized syllables closed with a geminate as well? We would assume that the answer is yes, but that is not the case in Fula and Koya, which are discussed in this chapter.

The answer offered here to the question of how languages vary with respect to having or not having hypercharacterized syllables and how languages can vary with respect to the type or types of hypercharacterized syllables they permit comes from two sources. The first part of the answer is representational. There are two ways for syllables to be hypercharacterized, being closed with a moraic consonant or with an appendix consonant, and there are two types of moraic consonants, simple and geminate. Neither of these premises is new or strange.

The second way in which languages may differ with respect to hypercharacterized syllables follows from Optimality Theory. All languages have *$\sigma_{\mu\mu\mu}$, but some have it ranked below other constraints that take precedence. There will be violations of *$\sigma_{\mu\mu\mu}$ in order to meet higher constraints. Other languages have *$\sigma_{\mu\mu\mu}$ ranked above the other relevant constraints, so *$\sigma_{\mu\mu\mu}$ is not violated and there are no tri-moraic syllables. The premises about structure in the previous paragraph are relevant, of course. As discussed in Chapter 2, a highly ranked *$\sigma_{\mu\mu\mu}$ is no bar to hypercharacterized syllables closed with an appendix consonant.

The purpose of this chapter is to describe languages which do not have a certain type of hypercharacterized syllable, a hypercharacterized syllable closed with a geminate, one type that

Finnish does have. We will see two languages, Fula and Koya. Fula is a West African language. Koya is a Dravidian language. Both lack hypercharacterized syllables closed with a geminate. Both do have hypercharacterized syllables closed with a simple consonant, which I analyze in both cases as appendixal. The desire here is to analyze languages which lack a certain type of hypercharacterized syllable with the same theory used to permit the same type of syllable in another language. Section 4.2 is devoted to Fula. Section 4.3 is devoted to Koya.

4.2 Fula

Fula is a west African language. My sources for Fula are Arnott's (1970) grammar and Paradis' (1992) dissertation, which contains new data, as well as a theoretic analysis. In many points, I will follow Paradis' analysis.

Fula has been of interest to phonologists for a variety of reasons, but we will concern ourselves mainly with the hypercharacterized syllables of Fula. Some discussion of the hardening of geminates is included in 4.2.4. Fula has morphemes which include empty prosodic material (an unassociated mora, in our terms). This mora associates to the final consonant of a stem, giving rise to geminates.

1)	stem	geminated form	gloss (Paradis p. 172)
	saw	cabb-i	stick
	lϵw	lϵbb-i	month
	wuy	gujj-O	thief
	lef	lepp-i	ribbon
	kOs	kOcc-ϵ	milk

Helping us to keep track of the geminates is another fact about Fula. The geminates of Fula are always non-continuant. Segments which are continuant underlyingly will, when they become geminates, map to a non-continuant surface structure as they do in the above examples. This is discussed in 4.2.4, below.

Fula, then, has geminates over morpheme boundaries, as well as in monomorphs. Fula lacks geminates after long vowels. In words where we would expect geminates to arise as a result of suffixes, we find instead a non-geminate consonant, but that consonant must be non-continuant. In other words, after a long vowel, where we would expect to find a geminate, instead we find a simple consonant and further that consonant must meet the constraint against [+continuant] that geminates must meet, though that is not required of simple consonants in any other environment of the language. Here are some examples of stems with a long vowel, with a morpheme which gives rise to gemination.

2)

stem	geminated form	gloss (Paradis p. 173)
Niiw	Niib-i	elephant
maay	maaj-i	river
lєєs	lєєc-є	bed

Fula has consonant clusters.[1] Although clusters occur over morpheme boundaries, the clusters also occur both at the end of roots and truly root internally, as shown below. Clusters may occur after long vowels, as below.

3) Consonant Sequences (Arnott p. 401, root underlined)

dept-e-re	book
^{m}*bort*-a	young sheep
kardumb-al	spindle
shirmball-al	twig[2]

4) Consonant Sequences after Long Vowels (Paradis p. 175)

kaakt-є	spittle
Naabl-u-dє	to yawn
fOOB-re	stump
caak-ri	couscous

Our task in analyzing Fula is to derive the hypercharacterized syllables closed by a sequence of consonants while at the same time account for the underparsing of the geminate's mora after a long vowel. This is accomplished by claiming that Fula has appendix consonants. The hypercharacterized syllables closed with consonant sequences are, then, bi-moraic and the constraint *$\sigma_{\mu\mu\mu}$ is unviolated. When an underlying sequence of a long vowel is followed by a geminate, it is more harmonic to leave the geminate mora unparsed, as opposed to violating *$\sigma_{\mu\mu\mu}$.

There is no other evidence that the consonant sequence of Fula are not moraic beyond the vowel length before a consonants sequence. A tri-moraic syllable closed with a consonant sequence can, in general, arise by a violation of *$\sigma_{\mu\mu\mu}$. Nor is there evidence that the consonant sequences are moraic. The evidence is better for the appendixal status of the consonants of Koya, which can be found later in this chapter.

---

[1]Fula also has pre-nasalized stops, which are identical to the clusters in the orthography of Fula. Care must be taken to avoid confusing the two when the status of clusters is at stake. I use no nasal-stop clusters in crucial examples below to avoid any confusion with the pre-nasalized stops.

[2]Probably: r^{m}b.

I will proceed with a list of the relevant constraints, all of which have been discussed earlier in this dissertation. Next, the constraints will be ranked. As in the discussion of Finnish, we will proceed with two candidates being evaluated by two constraints. Once we have gone through the available ranking arguments, those arguments will be summarized and full tableaux will be given.

4.2.1 The Relevant Constraints of Fula

In our discussion of Fula, we will need the following constraints.

5a)	$PARSE_{SEGMENT}$	Segments must be parsed.
b)	$*\sigma_{\mu\mu\mu}$	Syllables are maximally bi-moraic.
c)	$PARSE_{\mu}$	Morae must be parsed.
d)	$*\mu_{CONS}$	A mora may not be associated to a consonant.
e)	*APPENDIX	No appendixal consonant allowed.

These constraints have all been discussed above. In Chapter 2, they were used in a factorial typology. In Chapter 3, they were used to analyze Finnish, although in that case there were two sub-constraints for the one constraint *APPENDIX. In Fula, the two or more *APPENDIX constraints have no relevant constraint ranked between them, so they have the same effect one constraint against appendix consonants.

Some of the rankings of constraints is evident based on the description of Fula. Since Fula has geminates, $PARSE_{\mu}$ must dominate $*\mu_{CONS}$. Since Fula has consonant sequences, $PARSE_{SEGMENT}$ must dominate either $*\mu_{CONS}$ or *APPENDIX.

4.2.2 Ranking Arguments

This section has ranking arguments for this analysis of Fula. We will work with four underlying representations, given below.

6a)	/$le^{\mu}f^{\mu}i^{\mu}$/	"ribbon"
b)	/$ka^{\mu}rdu^{\mu}mbal$/	"spindle"
c)	/$Ni^{\mu\mu}w^{\mu}i$/	"to yawn"
d)	/$ka^{\mu\mu}ktE^{\mu}$/	"spittle"

The underlying form in (6a) has a short vowel followed by a geminate. The underlying representation in (6b) has a short vowel followed by a consonant sequence. Form (6c) has a long vowel followed by a geminate. Form (6d) has a long vowel followed by a consonant sequence. The underlying representations in (6a), (6b) and (6d) will be parsed faithfully. The form in (6c) will have an unparsed mora.

Our first ranking argument comes from $/le^{\mu}f^{\mu}i^{\mu}/$ "ribbon". This underlying form has a geminate after a short vowel. Such an underlying form can be used to rank the constraints $PARSE_{\mu}$ and $*\mu_{CONS}$.

7) Ranking Tableau for $PARSE_{\mu}$ and $*\mu_{CONS}$

		$/le^{\mu}f^{\mu}i^{\mu}/$	$PARSE_{\mu}$	$*\mu_{CONS}$
a)		{ l e_{μ} }{ $f_{<\mu>}$ i_{μ}}	*!	
b)	☞	{ l e_{μ} { f_{μ}} i_{μ}}		*

In this tableau, we see that candidate (7a) violates the constraint $PARSE_{\mu}$ while candidate (7b) meets that constraint. Candidate (7b) violates $*\mu_{CONS}$, which candidate (7a) meets. Because candidate (7b) is the correct form, the ranking in this tableau must be correct.

8a) $PARSE_{\mu}$ >> $*\mu_{CONS}$

Again, this ranking is necessary for a language to have geminates, as discussed in Chapter 2.

Our next ranking argument comes from two of the underlying representations given in this section. The first is the form $/ka^{\mu\mu}ktE^{\mu}/$, which has a consonant sequence as part of its underlying form.

This form allows us to rank $PARSE_{SEGMENT}$ and *APPENDIX. This is shown in the ranking tableau in (9), below.

9) Ranking Tableau for $PARSE_{SEGMENT}$ and *APPENDIX

		$/ka^{\mu\mu}ktE^{\mu}/$	$PARSE_{SEGMENT}$	*APPENDIX
a)		{ k $a_{\mu\mu}$ <k>} {t E_{μ}}	*!	
b)	☞	{ k $a_{\mu\mu}$ k} {t E_{μ}}		*

In this tableau, we see that candidate (9a) violates the constraint $PARSE_{SEGMENT}$ while candidate (9b) meets that constraint. Candidate (9b) violates *APPENDIX, which candidate (9a) meets. Because candidate (9b) is the correct form, the ranking in this tableau must be correct.

8b) PARSE$_{SEGMENT}$ >> *APPENDIX

These constraints are also rankable on the basis of the underlying form /ka$^{\mu}$rdu$^{\mu}$[mbal]/.

10) Ranking Tableau for PARSE$_{SEGMENT}$ and *APPENDIX

		/ka$^{\mu}$rdu$^{\mu}$[mbal]/	PARSE$_{SEGMENT}$	*APPENDIX
a)	☞	{ k a$_{\mu}$ r }{ d u$_{\mu}$ }		*
b)		{ k a$_{\mu}$ < r> }{ d u$_{\mu}$ }	*!	

In this tableau, we see that candidate (10b) violates the constraint *APPENDIX, while candidate (10a) meets that constraint. Candidate (10a) violates PARSE$_{SEGMENT}$, which candidate (10b) meets. Because candidate (10a) is the correct form, the ranking in this tableau must be correct.

This underlying form, /ka$^{\mu}$rdu$^{\mu}$[mbal]/, also permits us to rank the constraints *μ_{CONS} and *APPENDIX, as shown in the tableau below. In this tableau, we see that candidate (11a) violates the constraint *μ_{CONS}, while candidate (11b) meets that constraint. Candidate (11b) violates *APPENDIX, which candidate (11a) meets.

11) Ranking Tableau for *$\mu_{CO\ NS}$ and *APPENDIX

		/ka$^{\mu}$rdu$^{\mu}$[mbal]/	*μ_{CONS}	*APPENDIX
a)		{ k a$_{\mu}$ r$_{\mu}$ }{ d u$_{\mu}$ }	*!	
b)	☞	{ k a$_{\mu}$ r }{ d u$_{\mu}$ }		*

Because candidate (11b) is the correct form, the ranking in this tableau must be correct.

8c) *μ_{CONS} >> *APPENDIX

The underlying forms /ka$^{\mu\mu}$ktE$^{\mu}$/ and /ka$^{\mu}$rdu$^{\mu}$[mbal]/ both allow us to rank the constraints PARSE$_{SEGMENT}$ and *APPENDIX. The underlying form /ka$^{\mu\mu}$ktE$^{\mu}$/, however, does not permit us to duplicate the ranking argument between *μ_{CONS} and *APPENDIX. All of the candidates that have a violation of *μ_{CONS} will also have higher violations, of either *$\sigma_{\mu\mu\mu}$ or of PARSE$_{\mu}$. Consequently, those candidates would be ruled out whether *μ_{CONS} dominates *APPENDIX or not.

Finally, we turn to the underlying form that gives us our last ranking, /$Ni^{\mu\mu}w^{\mu}i$/. This underlying representation provides us with a ranking between the constraints *$\sigma_{\mu\mu\mu}$ and $PARSE_{\mu}$. In this tableau, we see that candidate (12b) violates the constraint *$\sigma_{\mu\mu\mu}$, while candidate (12a) meets that constraint. Candidate (12a) violates $PARSE_{\mu}$, which candidate (12b) meets.

12) Ranking Tableau for *$\sigma_{\mu\mu\mu}$ and $PARSE_{\mu}$

	/$Ni^{\mu\mu}w^{\mu}i$/	*$\sigma_{\mu\mu\mu}$	$PARSE_{\mu}$
a) ☞	{N $i_{\mu\mu}$ }{ $w_{<\mu>}$ i_{μ}}		*
	{N $i_{\mu\mu}$ {w_{μ}} i_{μ}}	*!	

Because candidate (12a) is the correct form, the ranking in this tableau must be correct.

8d) *$\sigma_{\mu\mu\mu}$ >> $PARSE_{\mu}$

As discussed in section 2.4, a geminate cannot be parsed as non-moraic. When a long vowel is present, conflict with *$\sigma_{\mu\mu\mu}$ is inevitable.

These are all the rankings that can be directly made from the forms discussed. The arguments from transitive rankings, where two constraints can be ranked based not on a direct ranking argument but rather on the ranking of other constraints, are included, with the summary of rankings in the next section.

4.2.2.1 Summary of Rankings

Below is a summary of rankings, given first in the form of a list, which contains the rankings and the form or forms on which they are based, as well as the number of the ranking tableau.

8a) $PARSE_{\mu}$ >> *μ_{CONS} /$le^{\mu}f^{\mu}i^{\mu}$/ (7)

b) $PARSE_{SEGMENT}$ >> *APPENDIX /$ka^{\mu}rdu^{\mu}$[mbal]/ (9), /$ka^{\mu\mu}ktE^{\mu}$/ (10)

c) *μ_{CONS} >> *APPENDIX /$ka^{\mu}rdu^{\mu}$[mbal]/ (11)

d) *$\sigma_{\mu\mu\mu}$ >> $PARSE_{\mu}$ /$Ni^{\mu\mu}w^{\mu}i$/ (12)

Below is a chart, based on the same rankings given above. The chart also makes clear the transitive rankings.

13) PARSE$_{SEGMENT}$ $*\sigma_{\mu\mu\mu}$

PARSE$_{\mu}$

$*\mu_{CONS}$

*APPENDIX

The constraint PARSE$_{SEGMENT}$ is not rankable with respect to any constraint except *APPENDIX, but I will put it at the top of the tableaux, equal to $*\sigma_{\mu\mu\mu}$.

4.2.2.2 Full Tableaux

Below are full tableaux for the four forms which we are discussing. Because the discussion in the ranking arguments section above was of necessity somewhat fragmentary, there will be some redundancy. As usual, the names of the constraints are abbreviated in the interest of space. Our first underlying representation, below, is that of a short vowel followed by geminate. There isn't very much to say about this form. It is the normal case where an underlyingly moraic consonant follows a short vowel and where that consonant's mora must be parsed. Because PARSE$_{\mu}$ dominates $*\mu_{CONS}$, this underlying form will map to candidate (14b), a form with a geminate, as in 2.3.2.

14) Full Tableau for Underlying Form /$le^{\mu}f^{\mu}i^{\mu}$/

	/$le^{\mu}f^{\mu}i^{\mu}$/	P_{SEG}	$*\sigma_{\mu\mu\mu}$	P_{μ}	$*\mu_{C}$	*AP
a)	{ l e_{μ}} { f $_{<\mu>}$ i_{μ}}			*!		
☞ b)	{ l e_{μ} { f_{μ}} i_{μ}}				*	

Below is an underlying representation with a short vowel followed by a consonant sequence. I have suppressed the structure of the irrelevant portion of the root, which is enclosed in square brackets. This form has push-pull interaction between $*\mu_{CONS}$ and *APPENDIX on the one hand and PARSE$_{SEGMENT}$ on the other. If PARSE$_{SEGMENT}$ is above one of the two other constraints in the ranking hierarchy as it is, then the language will permit consonant sequences, in this case, beginning with an appendixal consonant. Recall that a single *APPENDIX constraint is all that is necessary to describe this interaction, unlike the case of Finnish in chapter 3 and the Australian languages discussed in section 5.2, below, where more than one *APPENDIX constraint was necessary. Recall that this is simply a case where no relevant constraint separates the *APPENDIX constraints, making them, as a block, indistinguishable from a single constraint against appendix consonants.

15) Full Tableau for Underlying Form /ka$^{\mu}$rdu$^{\mu}$[mbal]/

	/ka$^{\mu}$rdu$^{\mu}$[mbal]/	P$_{SEG}$	*$\sigma_{\mu\mu\mu}$	P$_{\mu}$	*μ_{C}	*AP
a)	{ k a$_{\mu}$ r$_{\mu}$ }{ d u$_{\mu}$ }				*!	
b) ☞	{ k a$_{\mu}$ r }{ d u$_{\mu}$ }					*
c)	{ k a$_{\mu}$ <r> }{ d u$_{\mu}$ }	*!				

In the next tableau, we turn to a form with a long vowel followed by a geminate at underlying representation. This is the first time that *$\sigma_{\mu\mu\mu}$ becomes relevant in the discussion of the full tableaux. It should be noted that the mora that is associated to the consonant is not part of that consonant's lexical representation. That mora is, instead, added as part of some, but not other suffixes. As such, it is part of the underlying representations of those forms and subject to the same PARSE constraints as elements of the roots. Note that this is an *underlying geminate*, not a *derived geminate*, as discussed in section 2.3.5. That is, the mora is present at underlying representation and is, therefore, present in every candidate, even if unparsed.

16) Full Tableau for Underlying Form $/Ni^{\mu\mu}w^{\mu}i/$

	$/Ni^{\mu\mu}w^{\mu}i/$	P_{SEG}	$*\sigma_{\mu\mu\mu}$	P_{μ}	$*\mu_C$	*AP
a) ☞	{N $i_{\mu\mu}$ }{ $w_{<\mu>}$ i_{μ}}			*		
b)	{N $i_{\mu<\mu>}$ { w_{μ}} i_{μ}}			*	*!	
c)	{N $i_{\mu\mu}$ }{ w_{μ} i_{μ}}		*!		*	

This tableau is an example where a violation of $*\sigma_{\mu\mu\mu}$ proves fatal for a candidate. A syllable with a geminate closing an already heavy syllable doesn't make it in Fula. That the geminate will have its mora underparsed, not the vowel, reflects the prediction offered earlier in this chapter. It is always worse to parse a consonant's mora than it is to parse a vowel's mora (but see Koya below.)

Below we have an example with a long vowel followed by a consonant sequence. The constraint against tri-moraic syllables is still unviolated. The consonant sequences have an alternative to being parsed as moraic, unlike geminates, which are inherently moraic. Consonant sequences can be parsed with their first consonant as an appendix, as below.

17) Full Tableau for Underlying Form $/ka^{\mu\mu}ktE^{\mu}/$

	$/ka^{\mu\mu}ktE^{\mu}/$	P_{SEG}	$*\sigma_{\mu\mu\mu}$	P_{μ}	$*\mu_C$	*AP
a)	{ k $a_{\mu\mu}$ <k>} {t E_{μ}}	*!				
b)	{ k $a_{\mu\mu}$ <k_{μ}>} {t E_{μ}}	*!		*	*	
c)	{ k $a_{\mu<\mu>}$ k_{μ}} {t E_{μ}}				*!	
d)	{ k $a_{\mu\mu}$ k_{μ}} {t E_{μ}}		*!		*	
e) ☞	{ k $a_{\mu\mu}$ k} {t E_{μ}}					*

There are three candidates in the tableau that have the first member of the consonant sequence as moraic, (17b), (17c) and (17d). In (17b), neither the mora nor the consonant is parsed, giving rise to two violations. In (17c), the consonant is parsed as moraic and one of the vowel morae is unparsed, giving rise to violation of PARSE$_\mu$. The constraint *APPENDIX is low ranking. We have established above that *μ_{CONS} is ranked above *APPENDIX. In (17d), the consonant is parsed as moraic and both of the vowel's morae are parsed; the form is now in violation of the undominated *$\sigma_{\mu\mu\mu}$. Suppose that an appendix were not an option for this form for some reason and that we wished to posit (17d) as the winning candidate--it is phonetically identical to the surface form after all. For this to work, PARSE$_\mu$ would have to dominate *$\sigma_{\mu\mu\mu}$. If PARSE$_\mu$ dominated *$\sigma_{\mu\mu\mu}$, there would be no degemination of a geminate after a long vowel, as can be seen by referring back to the full tableau for the underlying form /N$i^{\mu\mu}w^{\mu}$i/.

4.2.3 Summary of Vowel Length Interactions

The language Fula has been analyzed here as having an undominated *$\sigma_{\mu\mu\mu}$, which accounts for the degemination of a geminate after a long vowel. The fact that the geminate is degeminated after a long vowel instead of shortening the vowel follows from the constraint against moraic consonants, which always makes it more costly to parse a consonant's mora than a vowel's mora (but see Koya in 4.3.2.3, below). Fula is also analyzed as a language with an appendix consonant. It is only through the means of an appendix consonant that the hypercharacterized syllables of Fula can be derived without also permitting the illicit geminate hypercharacterized syllable.

4.2.4 Fula Hardening

Recall the hardening of Fula geminates. The following alternations occur between geminates and simple consonants. The first column shows assorted continuants as simple segments. The next column shows what happens to them when they are geminates. The third column shows such a segment where the mora is not parsed in deference to *$\sigma_{\mu\mu\mu}$ (see 4.2.3.2).

18) Fula Hardening

simple	geminates	degeminated
w	bb	b
w	gg	g
f	pp	p
s	cc	c
h	kk	k

The two /w/'s must differ in underlying feature specification, as we can see by the fact that they map to geminates at different places of articulation.

In this section, I will discuss Fula Hardening, for two purposes. First of all, the hardening effects can be achieved within an Optimality Theoretic framework. Prince and Smolensky (1993, section 10.3.4) provide an analysis, which I sum up here. Second, a couple of facts do no follow from that account. I will propose how one will follow from *μ_{CONS} or NOCODA, from which *μ_{CONS} was developed.

The alternations in (18) raise questions, which I list below.

19i) When suffixing an empty mora, why does that mora surface as associated to a consonant rather than to a vowel?

ii) How can the hardening be analyzed?

iii) How can the hardening take place in forms where the surface consonant is not a geminate?

iv) Why does the geminate surface as simple in order to avoid a tri-moraic syllable, instead of the vowel surfacing as short, when such a syllable would also have only two morae?

Because these questions deal with the mapping of underlying to surface forms, they are questions that must be answered by any approach to Fula Hardening.

Paradis (1988) analyzes these alternations with repair strategies. The course of a derivation is an follows. First, suffixation takes place and the geminate is created.

20) Suffixation (adapted from Paradis 1988, (5))

X - X X	X - X X
C V	C V
[+cont]	

This geminate is in violation of a constraint, as below.

21) Constraint against Continuant Geminates (Paradis 1988, (2))

* X X

[+cont]

A repair rule changes [+cont] to [-cont]. If there is a tri-moraic syllable, another repair rule degeminates the geminate.[3] We may sum up the answers of the constraint and repair approach as below.

22i) (no answer)

ii) Hardening takes place to correct illicit form.

iii) Hardening happens before degemination.

iv) (no answer)

Paradis discusses possible answers to (19i) and (19iv). The answer to (19ii) is couched in terms of a repair strategy. Answer (19iii) is evidence derivation, with one repair following another.

Prince and Smolensky (1993, chapter 10) discuss Fula and propose the following analysis. There is a constraint against continuant geminates, just as in (21). A [-continuant] specification is possible at the expense of violating parse-feature. There is also a constraint against tri-moraic syllables. In addition to these there are Fill-x and Parse-x, both of which refer to the x slots used by Paradis. Morae will serve for these, however, for our purposes.

The Prince and Smolensky propose the constraint against continuant geminates is higher ranking than $\text{PARSE}_{\text{FEATURE}}$. Further, FILL_{μ} requires that the mora be filled. The constraint against [+cont] geminates is evaluated even when that mora is not parsed.

We may now revisit the questions discussed above. The questions in (18) are not limited to processional models and are stated in terms of the mapping of underlying forms to surface forms. Stating the answers to these questions in an Optimality Theoretic approach involves stating the constraint interactions that govern the selection of the optimal candidates.

23i) (no answer)

ii) This is the result of a violation of $\text{PARSE}_{\text{FEATURE}}$ to meet the more highly ranking constraint against continuant geminates.

iii) The constraint against continuant geminates is evaluated even if the mora (or x-slot) is not associated to a syllable.

---

[3] It isn't completely clear to me how this approach fits into the Theory of Constraints and Repair Strategies (Paradis 1988), since the segmental repair of the geminate takes precedence over the prosodic repair of the tri-moraic syllable, which is not the structure of the theory. This is not to say that the approach is not self-consistent and does not get the results that it claims.

iv) (no answer[4])

There are several possible ways to approach questions (19i) and (19iv); why the suffixed mora does not attach anywhere but the final consonant and why the geminate degeminates instead of having the vowel shorten. Paradis and Prince and Smolensky discuss several possibilities.

A possibility that Prince and Smolensky (1993) discuss, following Paradis, is that there is constraint against long vowels.[5] Paradis does not consider it to be a constraint, since it is not surface true. This constraint against long vowels would dominate *μ_{CONS}, as in the tableau below. If the ranking were the other way around, there would be no geminates in favor of attaching the geminate's mora to the vowel melodeme. This constraint plays a role in the analysis of Koya, later in this chapter and the analysis of Wiyot in 2.7.

24) Interaction of *LONG-V and *μ_{CONS} with Short Vowel

	/$le^{\mu}f^{\mu}i^{\mu}$/	*LONG-V	*μ_{CONS}
a) ☞	{ l e_{μ} { f_{μ}} i_{μ}}		*
b)	{ l $e_{\mu\mu}$ } { f i_{μ}}	*	

Unfortunately, this ranking of constraints chooses the wrong candidate in forms with a geminate after a long vowel, as below, where both candidates meet the undominated constraint *$\sigma_{\mu\mu\mu}$.

25) Interaction of *LONG-V and *μ_{CONS} with Long Vowel

	/$Ni^{\mu\mu}w^{\mu}i^{\mu}$/	*long-vowel	*μ_{CONS}
a)	{N $i_{\mu\mu}$ }{ $w_{<\mu>}$ i_{μ}}	*	
b) ☞	{N $i_{\mu<\mu>}$ }{ w_{μ} i_{μ}}		*

Perhaps this contradiction will disappear with additional constraints in a fuller treatment.

Turning to question (19iv), notice that if a geminate's mora is left unparsed in a candidate, that candidate does not incur a violation of *μ_{CONS}, while a candidate with a vowels mora left

[4]That is, Prince and Smolensky do not offer an answer in that section on Fula. As noted below, this prediction will follow from NOCODA, which was not discussed in that section.

[5]See Koya later in this chapter and Wiyot in 2.3.5.1.

unparsed still incurs the violation of that constraint. Therefore, all else equal, we would expect the geminate to degeminate, rather than the vowel to shorten. This analysis will follow equally well from NOCODA, but Prince and Smolensky do not discuss that constraint in their discussion of Fula.

This section has summarized analyses of Fula Hardening, those of Paradis (1992, 1988) and Prince and Smolensky (1993). The one novel element that I have to offer is to point out that $*\mu_{CONS}$ or NOCODA can have the effect of answering an as yet unanswered question in Fula, the question of why a candidate with an unparsed consonantal mora is preferable to one with an unparsed vocalic mora. As noted above, this does not answer the question of why a suffixed mora preferentially surfaces as associated to a consonant, not a vowel.

4.3 Koya

The purpose of this section is to show a language, Koya, where there are clearly appendix consonants. This language is analyzed as having a low-ranked *APPENDIX constraint, below both $PARSE_{SEGMENT}$ and $*\mu_{CONSONANT}$. We will see that all simple consonants of Koya are non-moraic, whether they follow a long vowel or a short vowel. This conclusion is reached partly on the basis of vowel length and partly
on the basis of an iambic shortening phenomenon, which is sensitive to geminates but not to non-geminate consonants.

Koya may also be a counter-example to the prediction above that geminates will always shorten if a mora is force to be left unparsed. Koya apparently shortens the vowel in this case. I provide both discussion of how this might be a vowel lengthening in syllables that are not closed with a geminate and an analysis with two separate $PARSE_{\mu}$ constraints, one that deals with vowel morae, one that deals with consonant morae. With either solution, the status of the first consonant in a sequence as appendixal is still clear and necessary.

4.3.1 Appendix Consonants in Koya

Koya[6] (Tyler 1969) is a Dravidian language, which permits intervocalic clusters, but no word initial or word final ones (p. 33-34). Below I include a chart, adapted from Tyler's chart (p. 34). I provide this chart to show that Koya has a wide variety of consonant sequences. Although there are some gaps in the list of imaginable clusters, all the simple consonants of Koya, except /h/, are represented in both first and second position in clusters. The first member of the consonant sequence is in the first column, while the second member is in the row across the top.

---

6 The consonant inventory of Koya is as below, from p. 6 of Tyler. I have altered his chart slightly, compressing "dental", "dental-alveolar" and "alveolar" into coronal and substituting /ß/ for /v/, according to the descriptions and transcriptions later in the chapter.

	labial	coronal	al-pal	palatal	velar	glottal
stops	p b	t d		ṭ ḍ	k g	
fric	ß	s				h
affric			č j			
nasal	m	n	ṇ			
flap		r				
lateral		l				
glide				y		

Table 4) Consonant Sequences in Koya

	p	t	d	ṭ	ḍ	k	s	ß	c	j	m	n	ṇ	l	r	y
p	✓	✓	✓					✓							✓	
b		✓						✓							✓	
t		✓	✓				✓	✓	✓			✓		✓	✓	✓
d		✓	✓					✓	✓	✓		✓		✓	✓	✓
ṭ		✓	✓	✓			✓	✓								
ḍ		✓	✓		✓			✓		✓						
k		✓				✓	✓								✓	✓
g		✓										✓				
s	✓	✓	✓			✓	✓	✓	✓		✓			✓		
ß		✓	✓			✓		✓				✓				
c		✓	✓			✓			✓							
j			✓			✓				✓						
m	✓	✓	✓			✓				✓	✓				✓	✓
n	✓	✓	✓	✓	✓	✓			✓	✓		✓			✓	✓
ṇ		✓	✓	✓	✓				✓	✓			✓			
l	✓	✓	✓			✓		✓			✓	✓		✓	✓	
r		✓	✓			✓	✓		✓			✓		✓	✓	
h		✓				✓		✓								
y		✓	✓			✓	✓	✓				✓				✓

The segments /b/ and /g/ where left off of the top row of this table for reasons of space. This omission caused the homorganic nasal-stop clusters ending in these stops to be omitted, as well as the geminate corresponding to these segments. There are clusters of rising, falling, and equal sonority. There is no general place requirement on the first consonant of the cluster; they are not required to, say, begin with a coronal or to have only one place node between the two consonants. In short, I assume here that the clusters of Koya are not subject to any special constraints on cluster occurrence.

The clusters of Koya are heterosyllabic according to Tyler's description. The second consonant of a cluster of two consonants is parsed as an onset. Analysis of Koya clusters, then, depends on parsing the first consonant of a sequence of consonants into the preceding syllable. To return to the diagram from the introduction, we are concerned with the licensing of C_1, below.

26) σ_1 σ_2

μ_1 μ_2

V_1 C_1 C_2 V_2

I am proposing that C_1 is an appendix consonant in Koya. This means that C_1 will be adjoined to the syllable node, with no intervening mora. The fact that Koya has clusters which begin with a non-moraic consonant is an effect of ranking *APPENDIX below the two constraints $PARSE_{SEGMENT}$ and $*\mu_{CONSONANT}$, as discussed in Chapter 2 (compare this with Chapter 3, where the appendix constraints are sensitive to place of articulation and separated by $PARSE_{SEGMENT}$).

That the consonant is parsed is clear from the fact that there are surface consonant sequences. I will show two pieces of evidence that the simple consonants of Koya are non-moraic in syllable final position. Showing that the consonants of Koya are non-moraic is made easier by the existence of geminates, which are clearly moraic, permitting a close comparison between long vowels and syllables closed by geminates on the one hand and syllables closed by non-geminates on the other. Before turning to this vowel length evidence, however, the stress system will be discussed.

4.3.2 Non-moraic Status of Consonants in Koya

In this section, data will be discussed that is relevant to the moraic status of consonants in Koya.

4.3.2.1 Stress

Stress is a common place to check for the moraic status of segments. The stress system of Koya is described as below.

"In a phrase strong stress regularly occurs on the first syllable, each following syllable being more lightly stressed, but still distinguishable from weak stress. Within a word stress occurs with long syllables, i.e., (C)VV, CVC, CVVC, CVVNC. Syllabic prominence occurs on vowels. Weak stress occurs with short syllables..."

The following examples are given.

27)	pikku	aaki	puungaari	taato
	kuud∂vaali	ginne	kaapuram	beske
	∂ndooru	l∂nje	vuyyaal	

This description might be taken as an argument that all consonants of Koya are moraic, not just geminates. On closer inspection, Tyler refers to two types of stress, phrase and word. All the words given have stress on the first syllable, except for the last one which has non-initial stress, despite having a geminate close the first syllable. We would expect to find stress on the first syllable of a two syllable word and these cases might generally be an effect of the phrasal stress rule. On the whole, the stress data seem hard to interpret and I will ignore these data, in favor of the vowel length data presented below, pending both better information on stress and a full Optimality Theory approach to the stress of the Dravidian languages. This point is taken up again below in 6.3.1 in the general context of the relationship between weight and length.

4.3.2.2 Iamb Conditioned Mora Underparsing

The first process of Koya that shows the status of the appendix consonant as non-moraic is an underparsing of a vowel mora conditioned by the iambic foot of the language. I will not provide a detailed analysis of this phenomenon. While the geminate inspired underparsing of a vowel mora discussed below is of importance to this dissertation, this iamb inspired underparsing is remote from our main concerns. It is important that geminates and non-geminates act differently. Geminates act like long vowels, as below, suggesting geminates, not simple consonants, are moraic after a short vowel.

A long vowel is mapped to a short vowel when it is followed by either a syllable with a long vowel or with a short vowel and a following geminate. That preceding long vowel surfaces as long when the following syllable is a syllable with a short vowel closed by a non-geminate consonant. Below, I include a formulation of the rule and Tyler's examples.[7]

28) V: → V / ___ CV: (see p. 39)
CVG

29) tung + a:ni + addu
tunganaddu "which does"

[7] We could, and I do, wish for more examples. Note also a couple of oddities. Vowels in initial syllables are not affected. Also, the last vowel in the second morpheme of the first example is lost by a general process, where the first of two adjacent vowels is lost. Also unexplained is why a geminate does not degeminate when it closes a syllable in the environment where vowels shorten.

ta:t + a:l + o:ru
ta:talo:ru "mother's fathers"

In other words, the following pattering exists between long open syllables, syllables closed with a geminate, and syllables closed with a non-geminate. Below, I show the syllables that trigger underparsing in (30a), where "G" stands for the first half of a geminate, and those that do not in (30b).

30a) Triggers underparsing

$VV]_{\sigma}$
$VG]_{\sigma}$

b) Fails to trigger underparsing

$VC]_{\sigma}$

A geminate is treated the same way as the second half of a long vowel by this phenomenon. We can begin to analyze this alternation as a mora being left unparsed to conform to the preferred iamb, given below.

31) $[\sigma \quad \sigma]_{Foot}$

$\mu \quad \mu\,\mu$

It can be seen, in an informal way, that the constraint enforcing the preferred iamb must dominate the constraint requiring parsing of a mora, $PARSE_{\mu}$, because a vowel mora is left unparsed to satisfy that constraint.

From this thumb-nail sketch of an analysis, we can see that we expect only bi-moraic syllables to trigger this mapping of the preceding vowel to a short vowel. The underparsing happens when the second heavy (bi-moraic) syllable is parsed as the head of the iamb. A heavy-heavy sequence is not an optimal iamb.

The point of describing this underparsing is showing the nature of the following syllable, which we might say, informally, triggers the underparsing. Underparsing takes place when the following syllable is closed with a geminate but not by a syllable closed with a non-geminate consonant. That environment only makes sense of geminates are moraic and other consonants are not moraic.

4.3.2.3 Vowels Length before Geminates

The next piece of evidence that simple consonants of Koya are non-moraic is the non-occurrence of long vowels before a geminate. Koya permits long vowels before consonant sequences, as below. Forms with more than one morpheme have the first morpheme in italics.

32)	le:ŋga	"calf"	p. 11
	pe:ŋku	"gods"	
	a:ṇḍ	"female"	p. 8
	pu:ŋgari	"flower"	p. 12
	gu:nji	"post"	
	go:tra	"lineage"	p. 115
	*da:l*gudda:n	"red cloths"	p. 121
	*ve:ns*o:matku	"while playing"	p. 121
	*gubba:l*tporro	"on hill"	p. 124

The situation can be summed up below, where "G" indicates the first half of a geminate and "C" stands for a non-geminate consonant.

33)	*	VVG]$_\sigma$
	ok	VVC]$_\sigma$

This distribution of long vowels can be transparently accounted for if geminates are and other consonants are not moraic. A long vowel is bi-moraic, as usual, and geminates are moraic, while other consonants are not. A long vowel followed by a geminate would creates a tri-moraic syllable, while a long vowel followed by a consonant does not.

In addition to this phonotactic generalization about the distribution of long vowels and geminates, there are cases where a long vowel is followed by a geminate as the result of concatenation. The result is a short vowel followed by a geminate. Zvelebil (1970) discusses this as a historical change to be found in the historical phonology of a variety of Dravidian languages. I give examples of the process in Koya below. Note that all but one of the examples in (34a) could be attributed to Iambic Underparsing. I will, however, assume Tyler is correct in his description.

34a) Long Vowels before geminates

a: + tt + a + payya	p. 38
[attapayya]	"after becoming"
o: + tt + o:ṇḍu	p. 38
[otto:ṇḍu]	"he brought"

ke: + tt + o:ṇḍu p. 39
[ketto:ṇḍu] "he told"

i: + tt + o:ṇḍu p. 39
[ittoṇḍu] "he gave"

b) Long vowels before clusters

tuŋg + ana: + n + ki p. 116
[tungana:ŋki] "for doing"

ko:y + si p. 127
[ko:ysi] "having cut"

4.3.3 Pre-geminate Underparsing

This section contains an analysis of Koya Vowel length alternations in terms of underparsing of a vowel's mora. This approach, then, is an analysis of Pre-Geminate Underparsing, which would be referred to as closed syllable shortening in process oriented terms. I will proceed by listing the constraints necessary for the analysis. The constraints will be ranked, again in pair-wise fashion, as in Finnish and Fula. The constraints will be summarized and put into a partial ranking. This ranking will be employed in full tableaux for the relevant candidates.

4.3.3.1 Constraints

I list all the constraints that are relevant for an analysis of Koya Pre-Geminate Underparsing below.

35a) PARSE$_{SEGMENT}$ Segments must be parsed.

b) $*\sigma_{\mu\mu\mu}$ Syllables are maximally bi-moraic.

c) PARSE$_{\mu}$ Consonantal morae must be parsed.

d) *LONG-V Long vowels are in admissible.

e) $*\mu_{CONSONANT}$ A mora may not be associated to a consonant.

f) *APPENDIX No appendixal consonant allowed.

Most of these constraints are familiar from preliminary discussion in the introduction and more extensive discussion in chapters 2 and 3. Here, again, there is only a single constraint dealing with appendixes. This is the same as in Fula, but different form the situation in Finnish. See chapter 5 for more discussion of appendix consonants.

Constraint (35d) was employed in 2.7. It was originally proposed in passing in Prince and Smolensky (1993, chapter 10, footnote 83), in discussion of Fula, following Paradis (1988). In Chapter 2, above, it served a rather different purpose, suggesting that constraint has several effects. To see why it is needed here, consider the following diagram. Candidate (36a) is disfavored from violating the highly ranked constraint against tri-moraic syllables. On this view, however, There are two equally good candidates that meet this constraint. Each leaves one mora unparsed. Either (36b) or (36c) could be the output. The candidates differ in no other way. As shown in this analysis in the next section, the use of *LONG-V can separate candidates like candidate (36b) and candidate (36c). This is essential, if we are to accurately predict the outcome in Koya, and how it is different from the outcome in Fula, above.

36) Interaction of $*\sigma_{\mu\mu\mu}$ and $PARSE_{\mu}$

	$/V^{\mu\mu}G^{\mu}V^{\mu}/$	$*\sigma_{\mu\mu\mu}$	$PARSE_{\mu}$
a)	$\{V_{\mu\mu}\{G_{\mu}\}\ V_{\mu}\}$	*!	
b)	$\{V_{\mu\mu}\}\{G_{<\mu>}\ V_{\mu}\}$		*
c)	$\{V_{\mu<\mu>}\{G_{\mu}\}\ V_{\mu}\}$		*

Another difference between the Koya case and the Finnish analysis (chapter 3) is that there is a single constraint on appendix consonants. In reality, this is no more than a case where $*APPENDIX_{LABIAL}$ and $*APPENDIX_{CORONAL}$ are not separated by any known, relevant constraint. This is the same as in Fula, above.

4.3.3.2 Ranking Arguments for Koya

Our first ranking argument comes from an example with a geminate after a short vowel, $/go^{\mu}b^{\mu}e^{\mu}/$. This argument concerns the constraints $PARSE_{\mu}$ and $*\mu_{CONS}$. These two constraints are necessary to predict the distribution of geminates from underlyingly moraic consonants. The following ranking argument will show that $PARSE_{\mu}$ will dominate $*\mu_{CONS}$. This ranking argument is completely parallel to the discussion of geminates in 2.3.2, as well as in Fula, above, and in Finnish in chapter 3, above. In that sense, it is not very new, but it is a fully consistent part of the analysis of Koya, offered here.

37) Ranking Argument for $PARSE_{\mu}$ and $*\mu_{CONS}$

		$PARSE_{\mu}$	$*\mu_{CONS}$
a) ☞	{ g o_{μ} { b_{μ} } e_{μ} }		*
b)	{ g o_{μ} } { $b_{<\mu>}$ e_{μ} }	*!	

In this tableau, we see that the first candidate violates the lower of the two constraints, $*\mu_{CONS}$, while the second candidate violates the higher of the two, $PARSE_{\mu}$. Because the first candidate is the actual form, we know that the given ranking relationship is the correct one.

38a) $PARSE_{\mu}$ >> $*\mu_{CONS}$

The next ranking argument comes from an underlying form with a short vowel and a consonant sequence. The underlying form is /$ba^{\mu}nda^{\mu}$/.

39) Ranking Tableau for $PARSE_{SEGMENT}$ and *APPENDIX

		$PARSE_{SEGMENT}$	*APPENDIX
☞	{ b a_{μ} n } { d a_{μ} }		*
	{ b a_{μ} <n> } { d a_{μ} }	*!	

In this tableau, we see that the first candidate violates the lower of the two constraints, *APPENDIX, while the second candidate violates the higher of the two, $PARSE_{SEGMENT}$. The two candidates differ in whether the first consonant of the sequence is parsed. Because the first candidate is the actual form, we know that the given ranking relationship is the correct one.

38b) $PARSE_{SEGMENT}$ >> *APPENDIX

The same form provides us with another argument about the ranking of *APPENDIX. Several pieces of evidence were presented that consonants, besides geminates, are non-moraic in Koya in section 4.3. This ranking argument relies on that analysis of consonants. Note however, that the stress data in section 4.3.2.1 was not entirely consistent with this idea about the status of the consonants. This point is returned to in section 6.3.

40) Ranking Tableau for *μ_{CONS} and *APPENDIX

		*μ_{CONS}	*APPENDIX
☞	{ b a_μ n } { d a_μ }		*
	{ b a_μ n_μ } { d a_μ }	*!	

In this tableau, we see that the first candidate violates *APPENDIX, the lower of the two constraints, while the second candidate violates *μ_{CONS}, the higher of the two. These two candidates differ in the structure of the first syllable. Because the first candidate is the actual form, we know that the given ranking relationship is the correct one.

38c) *μ_{CONS} >> *APPENDIX

In Koya, the constraint *LONG-V dis-prefers long vowels, but long vowels are permitted in the language (compare this with the analysis of Wiyot in section 2.3.5.1, where there were no underlyingly long vowels in the language and see also the discussion of Fula hardening earlier in this chapter.) This situation requires that *LONG-V be ranked below PARSE$_\mu$, as established below on the basis of the form /po$^{\mu\mu}$lu$^{\mu}$/, "rice powder".

41) Ranking Tableau for PARSE$_\mu$ and *LONG-V

		PARSE$_\mu$	*LONG-V
a) ☞	{ p $o_{\mu\mu}$ } { l u_μ }		*
b)	{ p $o_{\mu<\mu>}$ } { l u_μ }	*!	

In this tableau, we see that the first candidate violates *LONG-V, the lower of the two constraints, while the second candidate violates PARSE$_\mu$, the higher of the two. The two candidates differ in whether or not both the morae of the first vowel are parsed. Because the first candidate is the actual form, we know that the given ranking relationship is the correct one. Let me emphasize that the adoption of the constraint *LONG-V does not entail that a language, or all languages, will lack long vowels. It only entails that there is a pressure against long vowels.

38d) PARSE$_\mu$ >> *LONG-V

The last two ranking arguments come from an example with a long vowel followed by a geminate underlyingly. These ranking arguments establish the relationship of the constraint *$\sigma_{\mu\mu\mu}$ to the constraint PARSE$_\mu$ and of the constraint *LONG-V to the constraint *μ_{CONS}. They are established of the basis of /$o^{\mu\mu}t^{\mu}o^{\mu}$ṇḍu^{μ}/. Note that this is one way in which Koya differs from Fula, which was discussed earlier in this chapter. That is, in Fula, the geminate's mora was left unparsed in the event that parsing it would give rise to a tri-moraic syllable, under pressure of the constraint against tri-moraic syllables, while in Koya, a vowel's mora is left unparsed in the event that parsing it would lead to a tri-moraic syllable, under the pressure of the same constraint.

42) Ranking Tableau for *$\sigma_{\mu\mu\mu}$ and PARSE$_{\mu}$

		*$\sigma_{\mu\mu\mu}$	PARSE$_{\mu}$
a)	{ o $_{\mu\mu}$ { t$_{\mu}$ } o$_{\mu}$ }	*!	
b) ☞	{ o $_{\mu<\mu>}$ { t$_{\mu}$ } o$_{\mu}$ }		*

In this tableau, the second candidate violates the lower of the two constraints, PARSE$_{\mu}$, while the first candidate violates, *$\sigma_{\mu\mu\mu}$, the higher of the two. These candidates differ in whether the vowel's mora is parsed. Because the second candidate is the actual form, we know that the given ranking relationship is the correct one.

38e) *$\sigma_{\mu\mu\mu}$ >> PARSE$_{\mu}$

Finally, these two candidates differ in which of two vowel morae is left unparsed. The first candidate has a vowel mora left unparsed, which results in a violation of *μ_{CONS}, while the second has a geminate's mora unparsed, which results in a violation of *LONG-V. The two constraints are in conflict because the alternative to having a long vowel, in this case, is having a moraic consonant, so one or the other constraint must be violated.

42) Ranking Tableau for *LONG-V and *μ_{CONS}

		*LONG-V	*μ_{CONS}
a) ☞	{ o $_{\mu<\mu>}$ { t$_{\mu}$ } o$_{\mu}$ }		*
b)	{ o $_{\mu\mu}$ } { t$_{<\mu>}$ o$_{\mu}$ }	*!	

Because the first candidate is the actual form, we know that the given ranking relationship is the correct one.

38f) *LONG-V >> *μ_{CONS}

These are all the direct linking arguments available from these forms of Koya. These arguments are summarized in the next section.

4.3.3.2.1 Summary of Rankings

Below is a summary of the rankings established in the last section.

38a) $PARSE_{\mu}$ >> $*\mu_{CONS}$

b) $PARSE_{SEGMENT}$ >> *APPENDIX

c) $*\mu_{CONS}$ >> *APPENDIX

d) $PARSE_{\mu}$ >> *LONG-V

e) $*\sigma_{\mu\mu\mu}$ >> $PARSE_{\mu}$

f) *LONG-V >> $*\mu_{CONS}$

As can more easily be seen on the chart below, the rankings are completely established, except that $PARSE_{SEGMENT}$ cannot be conclusively ranked. A line indicates that two constraints are in an established ranking relationship, with the higher two being the higher ranked.

43) $*\sigma_{\mu\mu\mu}$ $PARSE_{SEGMENT}$

$PARSE_{\mu}$

*LONG-V

$*\mu_{CONS}$

*APPENDIX

Full Tableaux are given in the next section. Since $PARSE_{SEGMENT}$ cannot be completely ranked, I will include it as the highest of the constraints, set off from the others with a dashed line.

4.3.3.2.2 Full Tableaux

In light of the discussion so far in this chapter, I assume that the full tableaux are self-explanatory. I have abbreviated the names of the constraints, in order to fit them into the space provided.

44) Full Tableau for /go$^{\mu}$b$^{\mu}$e$^{\mu}$/

		P_{SEG}	$*\sigma_{\mu\mu\mu}$	P_{μ}	*L-V	$*\mu_{C}$	*AP
a) ☞	{ g o$_{\mu}$ { b$_{\mu}$ } e$_{\mu}$ }					*	
b)	{ g o$_{\mu}$ } { b$_{<\mu>}$ e$_{\mu}$ }			*!			

45) Full Tableau for /ba$^{\mu}$nda$^{\mu}$/

		P_{SEG}	$*\sigma_{\mu\mu\mu}$	P_{μ}	*L-V	$*\mu_{C}$	*AP
a)	{ b a$_{\mu}$ n$_{\mu}$ } { d a$_{\mu}$ }					*!	
b) ☞	{ b a$_{\mu}$ n } { d a$_{\mu}$ }						*
c)	{ b a$_{\mu}$ <n> } { d a$_{\mu}$ }	*!					

46) Full Tableau for /po$^{\mu\mu}$lu$^{\mu}$/

		P_{SEG}	$*\sigma_{\mu\mu\mu}$	P_{μ}	*L-V	$*\mu_{C}$	*AP
a) ☞	{ p o$_{\mu\mu}$ } { l u$_{\mu}$ }				*		
b)	{ p o$_{\mu<\mu>}$ } { l u$_{\mu}$ }			*!			

47) Full Tableau for /ko$^{\mu\mu}$ysi$^{\mu}$/

		P_{SEG}	$*\sigma_{\mu\mu\mu}$	P_{μ}	*L-V	$*\mu_{C}$	*AP
a)	{ k $o_{\mu\mu}$ y_{μ}}{s i_{μ}}		*!		*	*	
b)	{ k $o_{\mu\mu}$ <y>}{s i_{μ}}	*!			*		
c)	{ k $o_{\mu<\mu>}$ y_{μ}}{s i_{μ}}			*!		*	
d) ☞	{ k $o_{\mu\mu}$ y}{s i_{μ}}				*		*

48) Full Tableau for $/o^{\mu\mu}t^{\mu}o^{\mu}$ṇḍ$u^{\mu}/$

		P_{SEG}	$*\sigma_{\mu\mu\mu}$	P_{μ}	*L-V	$*\mu_{C}$	*AP
a)	{ o $_{\mu\mu}$ { t_{μ} } o_{μ} }		*!		*	*	
b) ☞	{ o $_{\mu<\mu>}$ { t_{μ} } o_{μ} }			*		*	
c)	{ o $_{\mu\mu}$ }{ $t_{<\mu>}$ o_{μ} }			*	*!		

The main point of this section has been to offer a formalization of the analysis which involves an appendix consonant--that is, a non-moraic coda. This has been accomplished by showing that geminates, which are inherently moraic, give rise to a underparsing of a vowel mora and that other consonants do not.

Koya is a language where, I have argued here, consonants are appendixal; they are associated to the syllable node at the end of the syllable with no intervening mora. I have shown two processes that, in combination, argue that simple consonants are appendixal, even while showing that the geminates of the language are associated to morae and are therefore not appendixal. Non-geminates after a long vowel do not give rise to underparsing, suggesting that simple consonants are not moraic. That non-geminate consonants do not have a mora after a short vowel is shown by the fact that syllables closed with a non-geminate consonant do not "trigger" an Iambic Vowel Mora Underparsing, while we expect any bi-moraic syllable to do so. Therefore, all simple consonants are non-moraic. This fact can be analyzed in Optimality Theory as the effect of a low ranked constraint against appendixes, *APPENDIX. We expect that other languages may have this constraint ranked above $PARSE_{SEGMENT}$ and $*\mu_{CONSONANT}$, and in such a case, the language would not have appendixes, in favor of moraic consonants or unparsed consonants.

Koya, then, shows us that we may profitably extend the notion of appendix from "extra-coronal position word-finally" to an association to the syllable node on a par with onset. Appendixes are predicted not to affect vowel length or footing and this prediction is borne out, even though other consonants (the geminates) do affect both vowel length and footing. Koya also exhibits behavior that is opposite of Finnish. Koya respects $*\sigma_{\mu\mu\mu}$ at the expense of PARSE_{μ}, while Finnish violates $*\sigma_{\mu\mu\mu}$.

4.4 Conclusions

Koya and Fula are two languages where the constraint against tri-moraic syllables, $*\sigma_{\mu\mu\mu}$ is in conflict with the constraint or constraints that require morae to be parsed. The two languages both require that an underlying form that would result in a [...VVG] hypercharacterized syllable under a faithful moraic parse map instead to a form with one mora left unparsed. The two languages enforce this differently, Koya parsing the geminate's mora at the expense of the vowel and Fula parsing the vowel's mora at the expense of the geminate. The analysis of Koya requires the constraint *LONG-V, which prohibits long vowels.

Sherer, Tim D. 1994 Prosodic Phonotactics. Available from GLSA Publications, Department of Linguistics, South College, University of Massachusetts, Amherst, MA 01003.

PREFACE TO THE COMPUTER-FILE VERSION

This is the fifth of six chapters of my dissertation. This version of my dissertation is divided into seven chapters, numbered 0 to 6. Chapters 1 through 6 contain the corresponding chapters of my dissertation. Chapter 0 contains the front matter, as well as the appendix and the sources cited.

This version of my dissertation was taken from the Graduate Linguistics Student Association (GLSA) version, which had several typos corrected over the deposit version. This version differs from the GLSA version with respect to some diagrams. Mainly in tableaux, I have substituted bracketing and subscripted morae for tree diagrams in representing syllables, as below:

$\{C\ V_{\mu}\}\ \{C\ V_{\mu}\}$	=	2 cv syllables
$\{C\ V_{\mu\mu}\}$	=	a syllable with a long vowel
$\{C\ V_{\mu}\ C_{\mu}\}$	=	a bi-moraic closed syllable
$\{C\ V_{\mu}\ C\}$	=	a mono-moraic closed syllable
$\{C\ V_{\mu}\ \{\ C_{\mu}\ \}\ V_{\mu}\ \}$	=	2 syllables with a geminate consonant

There are still some diagrams in their original forms. I hope that these are recoverable from context.

The pagination is different from the original version. Each chapter is paginated separately. There is a table of contents for each chapter at the beginning of that chapter.

This chapter focuses on the topic of syllable appendixes and appendixal consonants. The main part of the chapter is a discussion of several (closely related) Australian languages and why they should be analyzed employing appendix consonants. I provide a sketchy Optimality Theoretic analysis. The balance of the chapter is devoted to questions about the appendix position. The idea of analyzing appendixes by means of Alignment is discussed, including how this could derive different sites for adjunction (word, syllable, and so forth.) The nature of the coronal asymmetry is discussed and a phonetic scale is offered as a possible explanation.

Tim D. Sherer
TDSherer@Delphi.Com
December, 1994

APPENDIX CONSONANTS

CHAPTER 5

APPENDIX CONSONANTS

5.1 Introduction

Much of the discussion of chapters 2 through 4 has been concerned with the distribution of moraic consonants within individual languages and across languages. This has been done on the premise that it is possible to discover the distribution of consonants from the distribution of the morae to which they are attached. For instance, in a certain language, if a consonant must be moraic and if a mora cannot exist in a certain environment (say, as part of a tri-moraic syllable), then there will be no consonants in that environment (no tri-moraic syllables.) In exploring consonant distribution in this way, we also have to acknowledge the possibility of non-moraic consonants, which I am calling appendix consonants. Such consonants can exist without the benefit of a mora or without the benefit of a mora exclusively dominating it, depending on one's assumptions about the structure of the appendix. If the facts about the distribution of morae and the constraints on moraic consonants are to have any predictive power, we must also have a theory of when a consonant can exist without a mora.

I am adopting the term *appendix* for all non-moraic, non-onset consonants. The term has been used to describe syllable and word final "extra" consonants, as in work by Selkirk (1978), Steriade (1981), and Borowsky (1986, 1989).[1] I am expanding the term to cover those instances of syllable final consonants that are not moraic. Previously, these consonants had simply been described as non-moraic and their representation has been various.

This chapter serves three rather different purposes. The chapter does have a single overall goal. That goal is to explore the nature of the appendix and the range of *APPENDIX

[1]See also Hayes 1982, 1985 for consonant extrametricality.

constraints. This is vital to this dissertation because the analyses offered so far depend, in one way or another, on the appendix position. Consonant sequences that begin with non-moraic consonants must be analyzed if the factorial typology in chapter 2 is to be general. Finnish coronal asymmetries were analyzed in chapter 3 as a result of coronal asymmetries, which are, by hypothesis, restricted to appendix consonants. The hypercharacterized syllables of Fula and Koya have been analyzed structurally, by positing a difference between moraic and appendixal consonants. This chapter, therefore, is an examination of the appendix position.

The first of the three related topics here is to exploring some languages which have appendixal consonants. I will discuss three Australian languages. These three languages clearly have appendixal consonants at the ends of syllables word-medially. Evidence for this is adduced from stress, vowel length, and minimal word constraints. One of the attributes that I am claiming for the appendix consonant is preferences for coronal place of articulation. Australian languages are clear illustrations of this preference.

The second purpose is to explore the structure of the appendix. It has been assumed above that the consonant is adjoined directly to the syllable node at the end of the syllable. This representation explains why neither onsets nor appendix consonants are relevant for stress placement or vowel length. However, there are a number of other adjunction sites imaginable in the prosodic hierarchy. Some of these have been proposed in the literature as sites to which appendix consonants are adjoined. The appendix position will be analyzed in terms of Alignment constraints, following McCarthy and Prince (1993b).

Finally, the articulatory scale will be discussed. This is a scale of speed or quickness that was discussed in passing in Chapter 3. Different articulations are arrayed along this scale, with apical coronals as the fastest and labials and velars as the slowest. This scale serves as the phonetic basis of the coronal preference. By hypothesis, the coronal preference among codas is attributed here only to the appendix position, not to moraic consonants. This hypothesis can only be tested with more analyses of different languages.

There will be much more to say about non-moraic syllable-final consonants than I have to offer here. This chapter attempts to answer some questions and phrase other questions about appendixes.

5.2 Appendixal Consonants in Three Australian Languages

There are two differences between appendixal and moraic consonants. First, there is the mere fact that appendixal consonants lack morae. This one difference leads to the following three predictions. First, moraic consonants may interact with vowel length (as in closed syllable shortening) and Appendixal consonants will not. Second, moraic consonants are visible to stress rules *that count morae* while
appendixal consonant will be invisible to stress systems that count morae. Finally, moraic consonants will count for minimal word conditions (McCarthy and Prince 1986) that are stated in terms of morae, while appendixal consonants will not be relevant for those conditions. This is the same as the stress argument, but phrased in minimal word terms.

The second difference between moraic and appendixal consonants is that appendixal consonants may show a coronal asymmetry. That is, some languages may have only coronals as appendixes. This fact does not follow from the non-moraic nature of the consonants in any obvious way. This argument is made only on the basis of a preponderance of evidence. That is, there are several cases where consonants that are appendixal are also restricted to coronality. These include English, Dutch, Finnish (see chapter 3; Finnish also has only coronals in word final position, which is possibly a word-final appendix, see Karttunen (1971)), and a variety of Australian languages (see below).

Adopting a framework of violable constraints makes the work of figuring out the status of consonants more difficult. Consider the constraint $*\sigma_{\mu\mu\mu}$. In a system where this constraint is unviolated at some level of representation, we would simply need to look at vowel length in closed syllables to ascertain the moraic status of the closing consonant. If no hypercharacterized syllables were possible, then the consonant closing the syllable must be moraic. If hypercharacterized syllables are possible, then the consonant closing the syllable must be appendixal. This is true because such an approach assumes an unviolated $*\sigma_{\mu\mu\mu}$.

In a framework like Optimality Theory, where constraints are violated, the question is more subtle. If a language lacks Hypercharacterized syllables, then $*\sigma_{\mu\mu\mu}$ is unviolated and any syllable final consonant must be moraic. If there are Hypercharacterized syllables, then either $*\sigma_{\mu\mu\mu}$ is unviolated and there are appendixal consonants or $*\sigma_{\mu\mu\mu}$ is violated and there are moraic consonants.

In the discussion of Australian languages below, I adduce vowel length interactions, stress, and coronal preferences. Once the existence and properties of the appendix consonant have been demonstrated, however, and we are sure that we need such an element, using an appendix consonant in a language like Koya in Chapter 4 on the basis only of vowel length interactions is a more secure move.

5.2.1 Three Australian Languages

A number of native Australian languages have similar consonant sequences. As I discuss below, these languages permit two types of sequences. One type is a homorganic [+sonorant][-sonorant] sequence. The other type, which also involves [+sonorant][-sonorant] sequence, requires that the first consonant be apical and that the second differ in place from the first. I will analyze these as sequences with an appendixal consonant as their first member.

In this section, I will discuss the consonant sequences, stress, and hypercharacterized syllables or three Australian languages, Wargamay (Dixon 1979), Guugu Yimidhirr (Haviland 1979), and Gumbaynggir (Eades 1979). Because these three languages have quantity-sensitive stress systems, it is possible to be sure that stress is not affected by syllable-final consonants. Languages with non-quantity-sensitive stress would not do for this purpose, because we would not expect length or closed syllables to affect stress in such languages. Later in this section, I will offer a brief analysis of the consonant sequences of these languages.

5.2.2 Consonant Sequences in Australian Languages

Here I describe the consonant sequences of Gumbaynggir, Guugu Yimidhirr, and Wargamay. The sequences are homorganic or an apical followed by a non-apical. These facts are true in many Australian languages. There is some variation from language to language.

The consonant sequences of Gumbaynggir are listed below. These are the sequences found within a root. Capital "N" is a laminal nasal. Sequences in parenthesis have been found in one root.

1) Homorganic Sequences

mb	nd	nǥ	ŋg

2) Apical-initial Sequences

nb	nǥ	ng	nm	
lb	lǥ	lg		
(lm)	(lN)	(lw)		
rb	rǥ	rg	rm	rw
(yb)	yg	(ym)	(yw)	

3) Other Sequences

wg	(mŋ)	
(Nb)	(Ng)	(Nm)

Consonant sequences found over morpheme boundaries follow the same patterns. There are three sequences which begin with a laminal nasal, each attested in only one root. I have no explanation for these, but they are rare. There are two other consonant sequences in the list above that do not fit the generalizations, wg and mŋ. Although we can observe that both of these begin with a labial and end in a velar, I see no way to analyze these with the rest. I will disregard these in the discussion below.

The pattern is similar in Guugu Yimidhirr. The allowable sequences are described as "any homorganic nasal stop cluster" (p. 40), or one of the segments [l], [rr] (a digraph indicating an apico-alveolar), [r], [y], [n], or [nh] (a digraph indicating a lamino-dental) followed by a non-coronal stop or nasal or non-coronal nasal-stop sequence. As in Gumbaynggir, there are also some sequences with a lamino-dental
nasal as the first member, a fact for which I offer no explanation. I list the sequences described here below (<u>n</u> and <u>d</u> are lamino-dentals, n^r and d^r are apical retroflexes, n^y and d^y are lamino-palatals, and R is an apico-alveolar).

4) Homorganic

mb	nd	n^rd^r	<u>nd</u>	n^yd^y	ŋg

5) Apical initial

lb	Rb	rb	yb	nb
lg	Rg	rg	yg	ng
lm	Rm	rm	ym	nm
lŋ	Rŋ	rŋ	yŋ	nŋ
lmb	Rmb	rmb	ymb	nmb
lŋg	Rŋg	rŋg	yŋg	nŋg

6) Lamino-dental initial

n̲b	n̲g	n̲m	n̲ŋ	n̲mb	n̲ŋg

A difference between Guugu Yimidhirr and Gumbaynggir is three-consonant sequences are allowed in the former, provided they are apical-initial, the second and third members a homorganic sequence.

Wargamay has a very similar set of sequences. Homorganic nasal-stop sequences are allowed, as are [l], [r], [R] (a retroflex; my symbol), and [y] followed by a non-apical stop, nasal, nasal-stop sequence, or [w], as well as [n] followed by a non-apical stop or nasal. This differs from the other two languages in allowing sequences of an apical coronal followed by a non-apical coronal; Guugu Yimidhirr and Gumbaynggir allow only [+coronal][-coronal]. Wargamay resembles Guugu Yimidhirr in allowing three-consonant sequences, provided that the second and third consonants are a nasal-stop homorganic sequence. The sequences described are listed below.

7) Homorganic

mb	nd	n̲d̲	ŋg

8) Apical initial

lb	rb	Rb	yb	ld̲	rd̲	Rd̲	yd̲		
lg	rg	Rg	yd	lm	rm	Rm	ym		
ln̲	rn̲	Rn̲	yn̲	lŋ	rŋ	Rŋ	yŋ		
lmb	rmb	Rmb	ymb	ln̲d̲	rn̲d̲	Rn̲d̲	yn̲d̲		
ŋg	rŋg	Rŋg	yŋg	nb	nd̲	ng	nm	nn̲	nŋ

The symbol d̲ is a laminal stop, n̲ is a laminal nasal, r is an alveolar trill, and R is a retroflex fricative, flap, or short trill.

To sum up the facts about consonant sequences, all three of these languages allow both nasal-stop homorganic sequences and sequences beginning with an apical sonorant. The languages differ in details, as we would expect different languages (even closely related ones) to do. They share, however, these two types of sequences. Again, this fact is offered as support of a view that apical coronals are permitted in syllable final position in these languages. To see how linked clusters fit into this view, see 5.2.6

5.2.3 Stress

In Wargamay, a long vowel will receive stress. Long vowels occur only in initial position, although there is some length from stress on vowels in other positions. If there is no long vowel, stress is reckoned from the right edge of the word. Closed syllables do not attract stress. Below I include some examples of stress assignment in Wargamay.

9) Examples of Stress Assignment in Wargamay

Long Vowel:	mu:ba	stone fish
	gi:bara	fig tree
No Long Vowel:	bada	dog
	gagara	dillybag

Stress is not marked on any examples with consonant sequences. Dixon's description is explicit, however, in that long vowels are relevant for stress and no mention is made of consonant sequences.

In Gumbaynggir, stress will fall on a long vowel. If there is no long vowel, stress will fall on a syllable closed with a glide. Otherwise, the first vowel is stressed in bi-syllables and a non-final syllable receives stress in words of three or more syllables. Again, syllables with a closing consonant (except for a glide) do not attract stress.

Note that this system might be analyzed as not counting morae at all. Instead, it could be analyzed as strictly prominence based (Prince 1983, McCarthy and Prince 1986, Prince and Smolensky 1993). What is most critical for our point here is that there is no evidence from stress that consonants are moraic.

10) Examples of Stress Assignment in Gumbaynggir

Long Vowel:	mi:mi	mother
	ŋali:	"ldu inc, s, a"
No Long Vowel:	nami	woman
	naliwan	jewfish

The fact that a closing glide will make a syllable heavy for purposes of stress may well be a reflection of sonorant sensitive moraification. That is, the idea that morae may be assigned to consonants which are higher on the sonority hierarchy but not to those that are lower on the hierarchy (Prince 1983, McCarthy and Prince 1986, Zec 1988).

In Guugu Yimidhirr, long vowels require stress. In words without long vowels, stress is initial, with alternating secondary stress. Once again, closed syllables do not attract stress.

11) Examples of Stress Assignment in Guugu Yimidhirr

Long Vowel:	magiil	branch
	burraay	water
No Long Vowel:	nambal	stone
	bigibigi	pig

These three languages differ in their stress systems. However, in each case, there is no preference for stress on closed syllables, even though each language allows closed syllables. These stress facts suggest that the languages in question have non-moraic consonants.

5.2.4 Minimal Words in Guugu Yimidhirr and Gumbaynggir

Guugu Yimidhirr Gumbaynggir offer another argument that consonants are not moraic. Guugu Yimidhirr and Gumbaynggir, unlike Wargamay, have a small number of monosyllabic content words.[2]

Gumbaynggir has only twenty monosyllabic roots (Eades 1979, p. 264). With the exception of a particle, [waw] "also", and a few of the verbs, which are not stable for vowel length, these all have a long vowel.

Guugu Yimidhirr has a small number of monosyllabic roots. The content words all have long vowels, as noted by Haviland (1979): "except for a few particles all of these have long vowels and most are closed with a final consonant." (p. 40).

If we interpret this fact as a minimal word effect (McCarthy and Prince 1986), stated as a bi-moraic minimal word, then the consonants must not be moraic or we would expect monosyllabic words with a short vowel and a closing consonant.

5.2.5 Hypercharacterized Syllables

All three of these languages have Hypercharacterized syllables, as do many other Australian languages. Below, I offer examples from each language. I employ the orthography from the sources. The consonant sequences are between the first and second vowels. Length is shown as doubling of the vowel in question or with a colon, again, following the source.

---

[2]See Kroeger's (1989) analysis of Malay.

12) Hypercharacterized Syllables in Gumbaynggir
(Eades 1979, from p. 348-354)

bu:mbi	"break, smoke"
da:ndur	"sick, sickness"
da:ŋgi	"fingernail"
ŋa:mbul	"magpie"
bu:rgi	"body"
bu:rwa	"paint, melt"
bu:nmi	"catch, pounce on, fall on/in"
da:lga	"sing"

13) Hypercharacterized Syllables in Guugu Yimidhirr
(Haviland 1979, from p. 167-171)

dyuumbil	"swallow"
guunduu	"three"
maandii	"bring, take"
nguulbaan	"cloud"
waarmbal	"return"
yiilba	"share, split with each other"

14) Hypercharacterized Syllables in Wargamay
(Dixon 1979, from p. 17-20)

ga:nda	"to crawl"
bi:lbi:l	"peewee"

Again, the mere existence of hypercharacterized syllables in and of itself does not does not demonstrate that the language has non-moraic consonants. It is only one piece of information that is consistent with such an approach.

5.2.6 Approach to Australian Consonant Sequences

In this section, I will offer a sketch of an analysis of the consonant sequences of these Australian languages. This is not intended to be a complete analysis. For instance, I will not deal with the differences between the languages. The goal of this section is simply to show how the coronal condition on appendixes will permit us to have the two main types of consonant sequences, homorganic and apical-initial non-homorganic. In one sense, this is more ambitious than the earlier chapters of this dissertation, which do not attempt to derive the type of sequences possible with respect to the featural make-up of the consonants involved, only the existence of such sequences. On the other hand, such a move is necessary if we are to go beyond the rough characterization of possible consonant sequences offered in chapter 2.

I will employ the following constraints. These are not familiar from earlier in this dissertation, except that *APPENDIX constraints have been used in Finnish in chapter 5.

15a)	*APPENDIX$_{NONCOR}$	Non-coronal appendix consonants are ruled out.
b)	*APPENDIX$_{NONAPIC}$	Non-apical coronal appendix consonants are ruled out.
c)	*APPENDIX$_{APICAL}$	Apical appendix consonants are ruled out.
d)	*PLACE-PLACE	No two consonant place nodes may be adjacent.
e)	PARSE$_{FEATURE}$	Features must be parsed.

The first three are appendix constraints, derived from an extended phonetic hierarchy of the type discussed in section 5.5 below. They allow us to distinguish between the coronal sub-places. For another approach, see Prince and Smolensky (1993, chapter 9).

The next constraint is a prohibition against two adjacent [PLACE] nodes. The articulators under the place nodes need not be identical in order for these constraint to be violated. A linked cluster, however, will not violate this constraint because it has a single [PLACE] node. This specific one has been proposed by Yip (1991), although that analysis had other features not adopted here.

Finally, there is a constraint requiring that features be parsed. The only features that will be relevant to our discussion are the place features. Whether this should actually be stated as PARSE$_{PLACE}$ is an important, but here unanswered, question.

The constraints PARSE$_{SEGMENT}$ and *μ_{CONS} are undominated in this analysis. Furthermore, FILL$_{FEATURE}$ is also undominated, ruling out candidates with non-underlying place features.

We will consider, first, the following underlying representation: /amka/. This is a contrived form, which will be unfaithfully parsed. I have suppressed the morae of the vowels, as they are not relevant. A faithful parse of this candidate is not admissible in any of the languages under discussion. Below is a tableau with the five constraints under discussion in a ranking relationship.

16) Full Tableau for /amka/

	/a m k a/	*AP$_{NC}$	*AP$_{NAP}$	*PL-PL	P$_{FEAT}$	*AP$_{AP}$
a)	a m k a	*		*!		
b) ☞	a ŋ k a	*			*	

The first of these two candidates is a faithful parse, the second is a linked structure, where the nasal has had its place feature left unparsed. Other candidates, with a whole segment unparsed, are ruled out by the undominated PARSE$_{SEGMENT}$.

These two candidates both violate *APPENDIX-$_{NONCOR}$. In both cases, a non-coronal segment is adjoined to the syllable node. They tie and evaluation of the form is passed on to the next constraint.

The next constraint where a violation is incurred is *PLACE-PLACE. This constraint is violated by candidate (16a), but not (16b). Candidate (16b) does violate the next constraint, PARSE$_{FEATURE}$. The optimal form is (16b), with a linked structure. Note that this linked structure is in violation of the constraint *APPENDIX.

The next form considered has an apical coronal as the first member of a sequence. An example from Gumbaynggir is offered below.

17) gunbur "bored"

18) Full Tableau for /gunbur/

	/gunbur/	*AP$_{NC}$	*AP$_{AP}$	*PL-PL	P$_{FEAT}$	*AP$_{AP}$
a) ☞	gu n b ur			*		*
b)	gu m b ur	*!			*	

This tableau has two candidates parallel to the previous tableau. In this case, however, the constraint *APPENDIX$_{NONCOR}$ is not violated by the first of the two candidates, while it is violated by the second. The first candidate, with an apical-initial sequence, is possible.

I haven't offered ranking arguments here, but the ranking offered is fully motivated. The ranking *APPENDIX$_{NONCOR}$ >> *APPENDIX$_{NONAPIC}$ >> *APPENDIX$_{APICAL}$ is provided by the phonetic hierarchy on which these constraints are based (see chapter 3 and 5.5). The constraint *PLACE-PLACE must dominate PARSE$_{FEATURE}$, as we can see from the underlying form /gunbur/. Further, if PARSE$_{FEATURE}$ were above *APPENDIX$_{NONAPIC}$, then all coronals would begin non-homorganic sequences, not just apical ones.

The purpose of this example is to illustrate how the articulatory asymmetries of the appendix position can be used to derive certain consonant sequences. The same principle was employed in Finnish, in chapter 3, to a more limited range of data. If, as is hypothesized here, only appendixes, not moraic consonants, have these place asymmetries, then all cases of coronal preference could be analyzed in terms of the same appendix constraints.

5.3 Location of Appendix

An appendix consonant is one that is both syllable final and non-moraic. An appendix consonant is, in a very real sense, outside the syllable proper and irrelevant to the syllable's constituency. Such a consonant would have to be adjoined to some prosodic position to be licensed, in line with Prosodic Licensing (Itô 1986). If we could have a consonant adjoined at any level of the prosodic hierarchy, then there are several places where we could imagine a consonant could be adjoined. Below, I give several possibilities.

19) Possible Adjunction Sites

a)	b)	c)	d)
prwd	prwd	prwd	prwd
FT	FT	FT	FT
σ	σ	σ	σ
μ	μ	μ	μ
V C	V C	V C	V C

It has also been proposed that the appendix consonant is a separate constituent, such as a degenerate syllable (McCarthy and Prince 1990) or a lone word-final mora (Mccarthy and Prince 1986). In a similar vein, appendix consonants have been proposed to be onsets to syllables which have zero vowels (Burzio 1988, Kaye 1990).

Structure (19a) is in the spirit of Zec (1988) and Hayes (1989), where a consonant is associated to a mora, but not exclusively dominated by it (see also McCarthy and Prince (1986) for discussion of this structure). This structure respects the Strict Layering Hypothesis (Selkirk 1984), but see Itô and Mester (1992) and Selkirk (1993) for discussion of that point.

Structure (19b) has the appendix as an adjoined position to the syllable, rather as the mirror, in a structural sense, of the onset position. This is the basic approach taken in the earlier chapters, in order to account for phonotactics, where a consonant is non-moraic, but where sequences occur freely within words. The same sort of phonotactic effects will follow from (19a), of course. McCarthy (1979) proposes a similar structure, but with the consonant Chomsky-adjoined to the syllable. This difference is in part designed to derive stress facts of Arabic.

Structure (19c) has the appendix as adjoined to the foot. I know of no one who has proposed this.

Structure (19d) has the appendix as adjoined to the Prosodic word. This has been proposed by Rubach and Booij (1990), as well as by Borowsky (1986), in slightly different form. Borowsky has a word-final extrametrical consonant, as well as a word-final position restricted to coronals, to explain the asymmetries discussed by Fudge (1969) and Selkirk (1978).

While (19a) and (19b) make no different predictions that I am aware of, (19a) and (19b) together differ from the predictions of (19c) and (19d). Both (19c) and (19d) predict the distribution of appendixes with respect to elements above the syllable. That is, appendixal consonants will occur, according to these representations, only foot- or word-finally. As noted above, word-final appendixes have been proposed for a variety of languages.

This section has discussed the possibilities offered by analyzing appendixes as adjoined structures. The number of possible adjunction sites is perhaps too rich for the behavior of natural language. We are certain only of the existence of (19a) or (19b) and (19d). Finding a principled reason for the lack of (19c) and a way to choose between (19a) and (19b), for both the linguist and the language learner, would be a good thing.

The purpose of this section has been to try to understand the possible structures of appendixal consonants. This is important, on the one hand, in predicting where consonant sequences may occur, since a consonant sequence can arise with an appendixal consonant as its first member. On the other hand, the structures listed in (19) are a necessary precursor to understanding appendix consonants as the results of Alignment violations, which is discussed in the next section.

5.4 Appendix as an Alignment Violation

McCarthy and Prince (1993b) propose that the constraint NOCODA may be stated as an Align violation, as part of a generalized theory of Alignment. They do this with a constraint requiring the right edge of a syllable to line up with the left edge of a vowel. That constraint is as below.

20) Align (σ, R, V, R) (=McCarthy and Prince 1993b (38))

This constraint reads, "It is more harmonic to have the right edge of a syllable be aligned with the right edge of a vowel." McCarthy and Prince note that this constraint does not permit different argument settings and that this constraint is substantially fixed.

McCarthy and Prince also discuss the possibility of analyzing word-final appendix consonants and foot-final consonants in terms of alignment constraints (p. 17-18). To see how this works, look at diagrams (19b) through (19d). All of these cases have the property of not aligning the consonant with the syllable as a prosodic unit. To see what I mean by this, consider the prosodic hierarchy.

21) The Prosodic Hierarchy

PRWD

FT

σ

μ

The structure in (19b) has a consonant after the final mora of the syllable, as do all the rest. Structure (19c) has the consonant after the entire syllable. Structure (19d) has the consonant outside the syllable and the foot and, consequently, the end of the prosodic word does not coincide with the end of a syllable or foot.

Suppose we have a group of constraints that align the right edges of the various categories with each other. I give some of these below, assuming only constraints that relate constitutes in immediate domination.

22) Some Alignment Constraints

	Constraint	Structure Prohibited
i)	Align prwd, R, FT, R	adjunction to PRWD
ii)	Align FT, R, σ, R	adjunction to foot
iii)	Align σ, R, μ, R	adjunction to syllable
iv)	Align μ, R, segment, R	(none)

Structure (19b) is a violation of constraint (22iii), structure (19c) is a violation of (22ii), and structure (19d) is a violation of constraint (22i). If constraints (22i-iii) were in a sub-hierarchy as (22iii) >> (22ii) >> (22i), it would explain the ability of a language to have word final appendixes without other appendixes. In addition, we might also note that the use of appendix constraints permits us to more closely tie the appendix to other notions of prosody (as in Selkirk 1992).

5.5 Phonetic Scale

The notion of an appendix constraint was introduced in chapter 1, above, as one of the constraints that governs syllable constituency. The appendix constraint and the constraint against moraic consonants together did the work of NOCODA of Prince and Smolensky. In chapters 1, 2, and 4, *APPENDIX was exemplified as a single constraint against licensing a consonant as an adjoined position. A single constraint, however, has the problem of not being able to deal with the coronal asymmetries that played a role in chapter 3 and earlier in this chapter. The appendix constraint discussed above was blind to segment featural content.

My solution to this problem follows from a phonetic scale. Phonetic scales were used by

Prince and Smolensky as one element of Optimality Theory (see Prince and Smolensky 5.2 and chapters 8 nd 9). A phonetic scale arrays the segments of a language along a phonetic dimension. The inherent phonetic ranking given by that scale is the basis for partial rankings for the constraints which use that scale. Prince and Smolensky (1993, chapter 5, section 2, and chapter 8) discuss how sub-hierarchies of constraints can be derived by combining phonetic, markedness, relational hierarchies. For instance, Prince and Smolensky discuss the sonority hierarchy, a phonetic dimension along which all the sound of a language can be arrayed. Every single sound will have a place on this hierarchy, but no place is more harmonic than any other in any inherent way. Instead, there are a variety of other hierarchies, in particular, the peak-prominence hierarchy, which states that the peak of a syllable is more prominent than the margins of the syllable. We have, two hierarchies.

23) Two Scales

PEAK-PROM	peak > margin
Sonority	vowels > sonorants > obstruents

Prince and Smolensky employ the single ">" is used to denote a relation along a phonetic or markedness[3] dimension. These two scales will combine such that the most harmonic peak will be a vowel, the next most harmonic a sonorant, the next most harmonic (or least harmonic) an obstruent. Likewise, the most harmonic margin will be an obstruent, the next most harmonic a sonorant, the least harmonic, a vowel. This operation on the scales above gives us a pair of hierarchies of the harmony of segments in the relevant positions, as below.

24) Two Harmony Scales

peak-vowel →	peak-sonorant →	peak-obstruent
margin-vowel ←	margin-sonorant ←	margin-obstruent

These harmony scales can be converted into ranked constraints, as below.

25) Two Sub-hierarchies

*peak-vowel << *peak-sonorant << *peak-obstruent

*margin-vowel >> *margin-sonorant >> *margin-obstruent

Languages differ in how these constraints are interleaved and in where constraints like PARSE are ranked, but not in the relative rankings of the constraints themselves. That ranking comes from the phonetic scales, via the procedure of interaction of the scales.

The phonetic scale of place of articulation that I adopt arrays different places of

[3]Markedness in the absolute, not contextual sense.

articulation according to the speed of the articulation (see Kuehn and Moll (1976) for speed of articulation). The fastest segments are apical coronals, which involve the movement of a small mass (the tip of the tongue) a very short distance (the alveolar ridge or just behind the alveolar ridge). Next are other coronals, such as dentals and retroflex consonants, which involve moving more of the tongue either forward (in order to spread out along the teeth in the case of dentals) or backward (in order for
the top of the tongue to contact the hard palate in palatalized consonants). After other coronals, the bottom of the list has both velars and labials. I'm not sure why these articulations are slower, in the sense of calculations of mass and distance, but they have been found to be slower than the coronal articulations (see Kuehn and Moll, 1976). The whole phonetic scale is given below.

26) Phonetic Scale of Speed of Articulation

apicals > non-apical coronals > non-coronals

Prince and Smolensky (1993, chapter. 9) use a markedness scale like this one for delimiting the possible inventories of languages. That scale, however, did not include apicals as a separate category. If appendix consonants are violations of an alignment constraint, and if we consider the possibility that alignment of this type is sensitive to the amount of time that passes, through the means of the scale above, not through introduction of timing into the phonology, then apicals are the least violation, non-apical coronals are next worse, and non-coronals are the worst. Essentially, the scale in (22) is combined with the scale that says that the longer an appendix is, the worse it is. The result is the following ranked constraints.

27) The APPENDIX Sub-hierarchy

$*\text{APPENDIX}_{\text{NONCOR}}$ >> $*\text{APPENDIX}_{\text{CORONAL}}$ >> $*\text{APPENDIX}_{\text{APICAL}}$

The constraint $*\text{APPENDIX}_{\text{NONCOR}}$ rules out both labial and velar appendix consonants. I have found no evidence that there is an reason to separate the non-coronals in this hierarchy. The Partial ranking offered above not only conforms to the scale, but can be shown to be the correct ranking for Finnish, through transitive ranking by another constraint, $\text{PARSE}_{\text{SEGMENT}}$ (3.4.3.1).

Having apicals as a separate category facilitates the analysis of Australian languages offered above. There are, of course, alternative analyses. Prince and Smolensky (1993, chapter 9) consider an approach where apicals are pure coronals and non-apical coronals have two tokens of the feature [coronal], an extension of the feature geometric approach to multiply articulated segments (see Sagey 1986).

5.6 Conclusions

This chapter has dealt with certain aspects of appendix consonants as they are employed in this dissertation. It included a brief analysis of consonant sequences from Australian languages, showing how a sub-hierarchy of *APPENDIX constraints can derive those facts, including the possibility of linked structures in appendix position.

Also discussed was the possibility of different adjunction sites in the prosodic hierarchy, a theory which provides some results, but also seems to be a bit too rich, making predictions about foot-final appendixal consonants. This notion of adjoined appendixes goes hand-in-hand with treating *APPENDIX constraints as Alignment violations, also discussed above.

Finally, the phonetic scale of speed of articulation was discussed. This scale provides the raw materials for the *APPENDIX constraints, including the basis for ranking them as offered in this chapter and in chapter 3.

The topics covered in this chapter have been various. The solutions posed, if any, are tentative. This chapter is more an exercise in phrasing questions than answering them, which is the next necessary step. However, since the substantive assumptions that we make about the appendix will affect the other elements of the grammar, we will need to address the questions discussed here.

Sherer, Tim D. 1994 Prosodic Phonotactics. Available from GLSA Publications, Department of Linguistics, South College, University of Massachusetts, Amherst, MA 01003.

PREFACE TO THE COMPUTER-FILE VERSION

This is the sixth of six chapters of my dissertation. This version of my dissertation is divided into seven chapters, numbered 0 to 6. Chapters 1 through 6 contain the corresponding chapters of my dissertation. Chapter 0 contains the front matter, as well as the appendix and the sources cited.

This version of my dissertation was taken from the Graduate Linguistics Student Association (GLSA) version, which had several typos corrected over the deposit version. This version differs from the GLSA version with respect to some diagrams. Mainly in tableaux, I have substituted bracketing and subscripted morae for tree diagrams in representing syllables, as below:

{C V_μ} {C V_μ}	=	2 cv syllables
{C $V_{\mu\mu}$}	=	a syllable with a long vowel
{C V_μ C_μ}	=	a bi-moraic closed syllable
{C V_μ C}	=	a mono-moraic closed syllable
{C V_μ { C_μ } V_μ }	=	2 syllables with a geminate consonant

There are still some diagrams in their original forms. I hope that these are recoverable from context.

The pagination is different from the original version. Each chapter is paginated separately. There is a table of contents for each chapter at the beginning of that chapter.

This chapter deals with topics not dealt with earlier in the dissertation, including a comparison between the Optimality Theoretic approach and an approach where constraints are set to *on* or *off*. Also discussed is the Linked Gemination approach to the representation of geminates, which is at odds with the moraic theory of geminates. Another matter covered was the role of linking, both in connection with Linked Gemination and the Coda Condition Approach, including discussion of an approach to hypercharacterization that employs coda conditions.

Tim D. Sherer
TDSherer@Delphi.Com
December, 1994

OUTSTANDING QUESTIONS AND OTHER APPROACHES

CHAPTER 6

OUTSTANDING QUESTIONS AND OTHER APPROACHES

6.1 Introduction

The first five chapters of this dissertation have dealt with the licensing of consonants and vowel length. A number of related questions have not been dealt with because they fall outside that scope. Mainly, while the existence of consonants in certain environments has been predicted, the nature of the consonants has not. The coronal asymmetries attributed to the appendix did predict the nature of consonant sequences, as in Finnish (chapter 3), and in the Australian languages (section 5.2), and in both cases constraints dealing with the occurrence of place nodes or articulator nodes were proposed.

Many of the topics taken up in this dissertation, as well as related topics not taken up, have been addressed by other researchers. This chapter has two goals. One is to look at the nature of the consonant sequences and the specific consonants that make them up. The second is to compare the current approach with other approaches. These two goals are linked in that other approaches make predictions about the nature of consonant sequences and this approach should as well.

In discussing other approaches, a caveat is in order. This dissertation employs Optimality Theory. Its results often follow from Optimality Theory conceptions of the grammar, especially the typological predictions of Chapter 2. As discussed in the next section, some of the results of this dissertation can't be achieved without Optimality Theory. Consequently, the arguments offered here that this approach is to be preferred will often carry over to other Optimality Theory approaches, provided that those approaches have the same raw materials as this approach. For instance, the interactions of $*\mu_{CONS}$ and the PARSE constraints discussed in Chapter 2 will also work with the more general NOCODA of Prince and Smolensky (1993, 4.1 and following sections).

Below is a summary of some of the elements argued for in previous chapters. These principles deal with consonant sequences, geminates, and hypercharacterized syllables.

1) two types of coda consonants

   a) moraic consonants, which count for moraic content
      blind to articulator

   b) appendix consonants, which do not count for weight
      sensitive to place of articulation

2) moraic theory of geminates

3) tri-moraic syllable constraint

Taken together, these principles make a variety of predictions, as described in this dissertation. Other approaches to these predictions can differ in one of two ways. First, they can differ in not employing Optimality Theory. Second, they can differ in the actual content of (1) - (3); that is, theories can differ in structural assumptions.

The next section is devoted to comparing violable constraints to a more parametric approach. What I am calling a parametric approach is one where languages differ in having constraints turned *on* or *off*. That section will show that the results achieved in earlier chapters are genuinely linked to Optimality Theory. Using the same constraints in parametric fashion does not yield the same results without complications of other elements of the grammar.

Section 6.3 is devoted to questions about the representation of geminates, particularly representations that employ underlying linking, instead of an underlying mora, to mark geminates in the lexicon.

Section 6.4 is devoted to the Coda Condition Approach, which employs restrictions on codas to derive the distribution of consonant sequences. This approach can also embrace the linked representation of geminates taken up in 6.3, so these two sections are closely intertwined. The Coda Condition Approach makes predictions about the range of possible consonant sequences in a language.

Section 6.5 discusses the use of a the Coda Condition Approach as a theory of hypercharacterization. This approach was taken by Borowsky (1986)

Section 6.6 serves as a summary of the various points discussed in this chapter.

6.2 Constraints and Parameters

This dissertation employs a theory of constraint interaction, Optimality Theory. Suppose, instead, that constraints were parameters, set to either *on* or *off* in different languages, as previous work with constraints has assumed (see, for instance, Paradis 1988, Kaye 1990). On this view, languages differ in the constraints that they have as turned on, but the constraints themselves are not violated. The purpose of this section is to show that employing Optimality Theory captures generalizations about constraint satisfaction that parametric theories do not. Optimality Theory gets the result of having constraints be violated in optimal candidates by having other constraints ranked above them. This is not to say that Optimality Theory simply has a different way of parameterizing constraints, because ranking gives rise to a different result than parameterization. In particular, ranking permits cases where a constraint is violated in some cases and not violated in others in the same language, at the same level of representation. This point is made in different forms by Prince and Smolensky (1993) in various places.

Recall the example of consonant sequences in the Australian languages discussed in 5.2. The two tableaux are repeated below.

4) Full Tableau for /amka/

	/amka/	$*AP_{NC}$	$*AP_{NA}$	*PL-PL	$P_{FEATURE}$	$*AP_{AP}$
a)	a m k a	*		*!		
b) ☞	a ŋ k a	*			*	

5) Full tableau for /gunbur/

	/gunbur/	$*AP_{NC}$	$*AP_{NA}$	*PL-PL	$P_{FEATURE}$	$*AP_{AP}$
a) ☞	gunbur			*		*
b)	gumbur	*!			*	

In the first case, the constraint *PLACE-PLACE, which prohibits two adjacent consonantal place nodes, is the deciding factor between the two candidates. The first candidate violates it, the second does not. In the second case, the optimal form violates that constraint. In that case, the first candidate is already selected as optimal. The constraint *PLACE-PLACE is irrelevant in this case.

This is how a constraint interaction model can differ from a parameter model. The constraint *PLACE-PLACE is not set to *off* in this situation. That constraint rules crucially in the first example and is not relevant to the second example.

We might, however, decide to modify the constraint *PLACE-PLACE or use some theory of coronal specification in order to avoid this behavior. In other words, we might write that constraint in such a way that it would not be violated by the optimal candidate in the second example. Doing so, after all, obviates the need for the separate appendix constraints. The problem with this tack, however, is that the new version of *PLACE-PLACE will now be a more specific constraint, designed not to be violated by an apical-initial sequence, or that the apicals will have to have a different representation, such as underspecification of the coronal node, designed not to set off the *PLACE-PLACE constraint.

If we change *PLACE-PLACE to mention apicals specifically, that constraint loses much of its claim to universality. While there are clearly other cases where coronals are permitted in environments like this, those are not always restricted to apical coronals. In other words, in addition to *PLACE-PLACE being parameterized to *on* or *off*, it must be parameterized in its contents, a move which involves a loss in generality.

Adopting underspecification of the coronal node would permit *PLACE-PLACE to remain simple. However, if we adopt an underspecification of coronal node approach, as in Paradis and Prunet (1992), then there is another degree of freedom on which languages may vary. That is, whether all coronals are underspecified or whether only apical coronals are underspecified. The representation would have to differ from language to language. In one language, only apicals would be underspecified, in another, all coronals would be underspecified (see McCarthy and Taub 1992 for points like this.) Now the representation itself is parameterized, instead of the constraint *PLACE-PLACE.

Either approach can be avoided by separating out the special status of coronality and apicality into the appendix constraints and leaving *PLACE-PLACE alone. This is only possible if constraints are allowed to interact and to be violated. This example shows how the *APPENDIX constraints adopted here will have the desired results without parameterization and without adopting very specific constraints or representations that vary from language to language.

Another example of this type is to be found in Finnish, in Chapter 3. In that language, there were types of hypercharacterized syllables; those ending in a geminate and those ending in an appendix. Syllables with a long vowel followed by a moraic simple consonant were not possible. Recall the constraint $*\sigma_{\mu\mu\mu}$, which prohibits tri-moraic syllables. It is not possible for us to say that this constraint is *on* or *off* in Finnish. Instead, $*\sigma_{\mu\mu\mu}$ proves fatal to one type of tri-moraic syllable, while being violated by another type of tri-moraic syllable.

The options for capturing this same effect in a parametric theory are clear. The constraint $*\sigma_{\mu\mu\mu}$ could be made more specific to allow one type of hypercharacterized syllable, but not the other type, as in Paradis (1988, 1992) work on Fula, or other parts of the grammar would have to be complicated, such as moraification (perhaps making it so that a rule of Weight by Position could not apply after long vowels). However the analysis is accomplished in a parametric approach, the cost is a loss in generality. One element or another must be more specific.

Other examples of this type are possible, but these two serve to explain the nature of the contrast between Optimality Theory and a parametric approach. In order for parametric theories employing parametric constraints to work, the constraints must be specific or some structures must have representational escape-hatches. Optimality Theory offers simple constraints that interact consistently with each other to achieve the desired result.

6.3 Linked Gemination

One of the elements of this approach has been the use of the moraic theory of geminates. Adoption of this approach has permitted us to describe the distribution of geminates by using the interaction of PARSE_{μ} and $*\mu_{\text{CONS}}$ and further allows us to have a system that permits geminates and consonant sequences to cross-classify.

As noted above, the moraic theory of geminates involves the inclusion of a mora with a simple consonant at underlying representation. If the mora is incorporated into a syllable and the consonant is incorporated into the onset of a following syllable, the result is a consonant that is linked to two syllables by independent principles of the language in question.

Another approach to the representation of geminates takes the linking itself to be relevant. The underlying representation of a geminate is a consonant attached to two timing nodes. Here the linking is not derived but primary. This approach includes McCarthy (1979), Clements and Keyser (1981), and Levin (1985), and also Selkirk's (1990) Two Root Theory, which introduces feature geometry into the matter by employing root nodes as the timing units. Let us call these approaches,
collectively, *linked gemination*. There are some predictions made by linked gemination theories that are not addressed in this work. This is especially true of the Two Root Theory of geminates, which employs feature geometric elements in the representation of linked consonant sequences and of geminates, has room to make these predictions about the nature of geminates and consonant sequences based on that representation.

In this section, I will discuss two elements of linked gemination approaches. First, I will discuss the *Equal Weight for Codas* generalization of Tranel (1991), which is used to argue against the moraic theory of geminates. Next, I talk about some other predictions of the Two Root Theory.

6.3.1 Equal Weight for Codas

Tranel (1991) offers a critical view of the moraic theory of geminates, in favor of linked gemination. Tranel proposes a generalization, called *Equal Weight for Codas,* that all the codas of a language are equal with respect to weight. That is, if a simple consonant is moraic, then so will be a geminate consonant. The moraic theory of geminates is not consistent with Equal Weight for Codas, because it predicts that
geminates should always cause a syllable to be heavy for purposes of stress, while there is no such onus on other syllable final consonants. Since Equal Weight for Codas seems to be correct, moraic theory of geminates must be correspondingly inadequate.

One of the two languages that Tranel discusses as a counter-example to the idea that geminates will attract stress is Tübatulabal. Stress does clearly pass over geminates in Tübatulabal. The moraic theory of geminates is quite unable to handle this fact. There is, however, another fact, about that language. There are no geminates after long vowels (Voegelin 1935). This fact makes sense if geminates are moraic.

Take another case, the well-known case of stress assignment in Yupik (see Krauss 1985, Woodbury 1985, 1987, Hayes (manuscript)). Stress is assigned on either a long vowel or the second of two short vowels. Consonants, geminates and otherwise, are not relevant. Consider, however, the following facts.[1]

[1]A variety of interesting work has been done on the Yupik facts, including but not limited to Krauss (1985), Woodbury (1985), (1987), Hewitt (1992), and Hayes

6i) The second of two short vowels is lengthened if and only if it is in an open syllable.

ii) A certain vowel cannot lengthen. When it is in an environment to lengthen, instead the following consonant geminates.

iii) The first syllable of a sequence of two light syllables, if closed by a geminate, will have the geminate degeminate.

iv) Some Yupik languages have closed syllable shortening, referred to as *compression* because diphthongs are made short as well as simple vowels.

On the one hand, Equal Weight for Codas is certainly correct of these Yupik languages. The consonants, geminate or otherwise, are not relevant for stress assignment. On the other hand, whatever one does to analyze the facts in (6), one must refer to consonants interacting with vowel length and syllable weight.

Finally, recall the cases of Fula and Koya in chapter 4. The data from stress in these languages was absent or suggested that all consonants were moraic. On the other hand, neither permitted hypercharacterized syllables closed with a geminate, while permitting hypercharacterized syllables closed with a consonant sequence.

What is emerging, then, from the cases here and the cases that Tranel discusses, is that Equal Weight for Codas seems to be true. That is, Equal *Weight* for
codas seems to be true. That is, for purposes of footing, codas are all equivalent. At the same time, vowel length interactions can treat consonants as moraic even in languages where they are not relevant for stress purposes and, most especially, vowel length interactions can treat geminates differently from simple consonants. So codas seem to be uniform for metrical purposes, but not for length. Geminates might well be irrelevant for stress, perhaps along the lines of Hayes (Manuscript), but in such cases, there is still evidence that geminates are moraic.

6.3.2 Other elements of Linked Gemination

Because Linked Gemination proposes a representation for geminates that more closely resembles consonant sequences, having two timing nodes, there is the potential in such a theory for analyses where clusters and geminates are related, either in distribution or by phonological rule. This is particularly true of the Two Root Theory (Selkirk 1990), where geminates are a subset of homorganic clusters. In this section, I will discuss some of these predictions. I do not have alternative analyses for all these facts.

One way in which a Linked Gemination theory can differ from the moraic theory of

---

(manuscript). Discussion of these works and the points rased therein is beyond the scope of this dissertation.

geminates is to predict parallels between the distribution of consonant sequences, in particular, linked consonant sequences, and geminates. The logic is that if a language permits a linked structure in an environment, we would expect it to permit other linked structures. Therefore, we predict that there should be languages with geminates and linked sequences in parallel distribution.

There are certainly languages that fit this description. Japanese and Diola Fogny (Sapir 1965) are two well known examples. The problem with this line of reasoning is that there are also languages which have linked consonant sequences and lack geminates, like English, and also languages with geminates that lack any consonant sequences, like Woleaian. These differences, then, must follow from different principles of the grammar. If this is true, then we need to ask if the different principles of the grammar could also account for the languages, like Japanese or Diola Fogny, that have only linked sequences and geminates without drawing on the structural similarity offered by Linked Gemination. The system offered in Chapter 2 predicts a cross-classification of consonant sequences and geminates. It lacks the means to distinguish between linked and non-linked sequences, but this failure could be remedied along the lines offered in 5.2 or in the discussion below. I will return to this point below.

A second type of prediction offered by Linked Gemination is the phonological relations between geminates and consonant sequences. That is, it is possible to derive one from the other by phonological rule. If geminates are linked structures with a structure similar to linked sequences, processes that change one to the other are not difficult to work out. On the other hand, if we adopt the moraic theory of geminates, creating a consonant sequence from a geminate is a complicated process, involving the insertion of a new segment.

A case of this can be found in Icelandic (Thrainsson 1978), where geminates will break into a sequence of /h/ followed by a simple version of the former geminate. This is referred to as Preaspiration. Selkirk (1990) refers to this type of phenomenon as *Laryngeal Fission*, since it is a case where the laryngeal features are separated into two separate laryngeal specifications on the two halves of the geminate.

Selkirk analyzes preaspiration as a simple delinking. Since the Two Root Theory developed in Selkirk (1990) has representations with two root nodes, there is no need to insert a separate segment. The only change that has to be made is to cut the association line between the root node and the shared laryngeal node, to take a laryngeal specification from the preceding vowel. This approach follows the general outline of Thrainsson (1978). Selkirk also analyzes facts from Klamath along these lines.

Two Root Theory also provides an analysis for diphthongization. In moraic theory, a long vowel has two morae. If it is diphthongized, the vowel melody must split. In the Two Root Theory, a long vowel has two root nodes and can, therefore, have the second root node's connection to the vowel melody severed (see Hayes (1990) for a co-indexation approach to these phenomena).

Finally, Selkirk offers an approach to geminate inalterability based on feature dependency. Since some features are dependant on others in the feature geometry, some of the features will have to be shared between the two root nodes. The dependant features (place features on stricture features) must be dependent on identical heads, by Selkirk's Homogeneous

Stricture Linking. Therefore, no rule may alter one of the higher nodes, under the two root nodes, without altering them both--inalterability. See Inkelas and Cho (1993) for another approach to inalterability and discussion of the Two Root Theory.

I have no analysis for Icelandic Preaspiration or other cases of geminate breaking or diphthongization. There is the possibility, of course, that these cases are at a separate level of representation. That is, Linked Gemination would be true after a linearization, perhaps along lines suggested in McCarthy (1986). This would explain why breaking respects moraic structure. Or, alternatively, it is possible in Optimality Theory that whatever constraints militate against Linked Gemination are violated in favor of other constraints in certain circumstances.

Either approach would help address the question of geminates derived from sequences. In the system offered in Chapters 2 through 4 of this dissertation, there is no way for a geminate to be derived from an underlying consonant sequence. This was a flaw in the analysis of Italian offered in 2.6. If it is more harmonic not to have a consonant sequence, then a PARSE or FILL violation will be more harmonic, since the first consonant of an underlying sequence has no mora to be preserved. In the Two Root Theory, this gemination would be preferable, since it permits a root node to be parsed, as apposed to being left unparsed, which must be more harmonic, provided geminates are permitted in the language.

It is fair to ask, then, if the moraic theory of geminates should not be replaced with a theory of Linked Gemination. This brings us back to the point of Chapter 2, in particular, that geminates and consonant sequences *cross-classify*. Cross-classification works in the moraic theory of geminates because geminates and consonant sequences have very dissimilar underlying representations. In a Linked Gemination approach, they have very similar underlying representations. This is not to say that a Linked Gemination approach could not capture cross-classification, only that I can't find a way to do it.

6.4 Linking and The Coda Condition Approach

Much of this dissertation has been devoted to deriving the existence or non-existence of consonant sequences from the possibility or impossibility of syllable final consonants in a language. This section is devoted to a theory of syllable-final consonants, the Coda Condition Approach (Itô 1986, 1989, Yip 1991, 1993, Zec 1993), which employs featural restrictions on codas. Further, that approach gives linked structures a special status, exempting them from coda condition.

Below, I give an example of a geminate and a consonant sequence from Finnish.

7a) laitta b) kulta

The geminate, to be found in (7a), is marked with a mora in the lexicon. Its other properties, such as ambisyllabicity, follow from this marking. The consonant sequence in (7b) on the other hand, is made up of two simple consonants, neither of which is marked as moraic. Whether a sequence surfaces depends on whether its members can be licensed. The consonant /t/ will be licensed as an onset in any theory. Licensing of /l/ can be achieved by parsing /l/ as moraic (as in Finnish) or as appendixal (as in Fula, or Koya (Chapter 4)). If it cannot be so parsed, that consonant will not surface and there will be no sequence.

Below, I give the representations of a geminate, a homorganic cluster, and a non-homorganic cluster in the Two-root Theory of Geminates (Selkirk 1990), which was discussed above. For simplicity, I omit the feature [consonantal] from the root nodes.

8) geminate [-son] [-son]

k

9) linked cluster [+son] [-son]

[+nasal] k

10) other cluster [+son] [-son]

l p

I've chosen the two root theory here because it makes clear even in these simple diagrams the relationship between clusters and geminates in that theory. Geminates and homorganic clusters are both linked structures. Linked clusters are geminates where certain features are not identical between the two halves of the "geminate".

The theory predicts that consonant clusters and geminates will act alike with respect to the phonological processes that take into account linking, a point mentioned in the previous section. Itô (1986) makes use of this principle of relating linking and licensing, in proposing coda conditions. We may refer to this as the Coda Condition Approach. I will discuss it as an approach where linking is used to describe the distribution of geminates and consonant sequences.

In Japanese, Itô proposes a coda condition, which is a restriction on the type of material that may occupy the coda of a syllable, or the second mora of a syllable in our terms or in the terms of Itô (1989). I give the formulation from Itô (1989) below.

11) * μ

[PLACE]

This constraint says that no consonant with a place node may occupy a mora in Japanese. Japanese has nasal-stop clusters, but they must be linked sequence. The Coda Condition Approach employs the Linking Condition (Hayes 1985) which requires structural descriptions be interpreted as exhaustive. The constraint in (11) will not apply to linked structures because (11) does not mention linking.

Itô (1989) employs the moraic theory of geminates. Itô (1986) does not use that approach, rather a linked geminate approach. The Two Root Approach can also take advantage of the Linking Condition to describe the licensing of geminates.

This system makes predictions about the set of clusters that a language may have, which has not been an element of the current approach so far. Linked structures will be favored in such an analysis. Further, other generalizations with respect to clusters will be possible, based on the range of possible coda conditions.

We need to ask two question about the coda condition approach. First, what range of geminates and consonant sequences does it predict? Second, how does it predict that a language fails to have sequences or geminates?

The answer to the first question is that the coda condition system predicts that clusters that will exist in a languages are those where C_1 of C_1 C_2 meets the coda condition of a language and linked sequences. In answer to our first question, so far, so good. There seems to be a preference for linked structures in the world's languages, a fact captured by such approaches, and not by the approach given in the first four chapters of this dissertation. Further, some languages have restrictions on the first member of a sequence, a fact captured in the current approach only by the place asymmetries entailed by appendixal consonants.

The answer to our second question is not so clear. We can see how this approach can *not* create non-linked sequences. The coda condition does not permit any consonant. How, though, does it prohibit *linked* structures? Linked clusters are
not subject to coda conditions and neither are geminates in a non-moraic theory of geminates. Any language should allow linked structures. The fact that a language lacks linked clusters must follow from the fact that linked clusters must not be available in such languages in the first place at all. Recall Woleaian (Sohn 1975), which has been discussed above (chapter 2, sections 2.3.2 and 2.3.3, and in passing). Woleaian lacks consonant sequences, except in recent loan words from English and Japanese. The language has geminates. In other words, it resembles Japanese and Diola Fogny, except for lacking consonant sequences.

Assuming the Moraic Theory of Geminates, that fact that the language has geminates follows from the ability of segments to be dominated by a mora and serve as an onset. In our terms, the constraint PARSE_{μ} dominates $*\mu_{\text{CONS}}$. The fact that there are no consonant sequences is captured in the current theory by not allowing any consonant to be licensed syllable finally.

Consider this from a Coda Condition point of view. How can we prohibit consonant sequences? Prohibiting all consonants in coda position will not work. That will prohibit simple consonants in coda position. A linked structure, however, will not be subject to that constraint by the Linking Condition, as mentioned above.

If we adopt a Linked Gemination approach, the problem is compounded. Then we must figure out how geminates, but not homorganic sequences, are permitted in that language.

The way that linked sequences can be prohibited in a Coda Condition approach is by prohibiting the rule of assimilation that creates them. In non-process oriented terms, we could prohibit linked structures. This done, there is no work for the Coda Condition to do. In other words, in an Optimality Theoretic approach, the Coda Condition Approach can be adopted directly without the use of coda conditions themselves.

As for the approach taken in this dissertation, it is modular in the sense of Lamontagne (1993). That is, the prosodic effects discussed in chapter 2 are separated from the featural effects discussed briefly in chapter 3 and again in 5.2. On this view, linking comes from the interaction of features (assimilation and dissimilation of features, in non-process-oriented terminology). The licensing of consonants is strictly a prosodic affair, derived from the interaction of PARSE, FILL, *μ_{CONS}, *APPENDIX, and so forth.

6.5 Hypercharacterized Syllables from a Coda Condition Approach

The Coda Condition Approach has been used as a theory of hypercharacterization. Borowsky (1986) notes examples of the following type in English, all of which have a long vowel followed by a consonant sequence.

12) Hypercharacterized Syllables in English

chamber, angel, ancient, danger, maintain, council, wainscot, cambric, dainty, laundry, Cambridge, boulder, shoulder, cauldron[2]

Borowsky notes that such forms all have a homorganic cluster and analyzes these forms with a coda condition. The coda condition is given below (this refers to the non-onset portion of the syllable, despite the x notation).

13) * x x x]$_\sigma$

C

In other words, the third element of a syllable's nucleus must not be a consonant. This analysis follows Itô's (1986) approach to Diola Fogny. The forms in (12) evade this condition by having linked structures. That is, the long vowel in a closed syllable is permitted due to the linking.

If we were to try to apply this coda condition analysis to other languages, there would be serious difficulties. First, recall the patterns of Fula and Koya. Neither language permits a geminate after a long vowel. Both languages permit consonant sequences, both homorganic and non-homorganic, after both long an short vowels. This will not follow from the template in (13), which would permit any linked structure to close a hypercharacterized syllable. Further, Finnish lacked certain linked clusters after long vowels, while permitting geminates freely. This fact, as well, does not follow from a template like (13).

To sum up this subsection, the coda condition approach gets some of the linking facts about the distribution of possible clusters, but cannot keep a language from having any clusters at all, nor keep a language from having geminates. The use of linking seems to over-predict hypercharacterized syllables being closed with linked structures. The current approach for the linking facts is that they follow from segmental considerations. The current approach to geminates is that languages have or don't have them based on the ranking of PARSE$_\mu$ and

---

[2] There are also the following examples form Borowsky: deictic, deixis, seismic. These are not accounted for under the coda condition, but have another explanation.

$*\mu_{CONS}$. The constraint $*\sigma_{\mu\mu\mu}$ is independent of linking and consequently, hypercharacterized syllables are not dependent on linking.

6.6 Summary

This chapter has dealt with topics that were not dealt with in the earlier parts of the dissertation. These include a comparison between the Optimality Theoretic approach taken here and a parameterized approach, where constraints are set to *on* or *off*. Also covered was the Linked Gemination approach to the representation of geminates, which is at odds with the moraic theory of geminates. Another matter covered was the role of linking, both in connection with Linked Gemination and the Coda Condition Approach, including discussion of an approach to hypercharacterization which employs coda conditions.